AD-A116 610 MARYLAND UNIV COLLEGE PARK INST FOR PHYSICAL SCIENCE--ETC F/G 12/1
THE FINITE ELEMENT METHOD FOR PARABOLIC EQUATIONS. I. A POSTERI--ETC(U)
APR 82 M BIETERMAN, I BABUSKA N00014-77-C-0623
UNCLASSIFIED BN-983 NL

1 OF 1
AD A
116610

END
DATE
FILMED
08-82
DTIC

INSTITUTE FOR PHYSICAL SCIENCE
AND TECHNOLOGY

15

AD A116610

Laboratory for Numerical Analysis

Technical Note BN-983

# THE FINITE ELEMENT METHOD FOR PARABOLIC EQUATIONS

## I. A POSTERIORI ERROR ESTIMATION

by


M. Bieterman

and

I. Babuška






April 1982



UNIVERSITY OF MARYLAND

ENCL (1)

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER<br>Technical Note BN-983 | 2. GOVT ACCESSION NO.<br>AD-A116 | 3. RECIPIENT'S CATALOG NUMBER<br>610 |
| 4. TITLE (and Subtitle)<br>THE FINITE ELEMENT METHOD FOR PARABOLIC EQUATIONS<br>I. A POSTERIORI ERROR ESTIMATION | | 5. TYPE OF REPORT & PERIOD COVERED<br>Final-life of the contract |
| | | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s)<br>M. Bieterman and I. Babuška | | 8. CONTRACT OR GRANT NUMBER(s)<br>ONR N00014-77-C-0623 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Inst. for Physical Science & Technology<br>University of Maryland<br>College Park, MD 20742 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Department of the Navy<br>Office of Naval Research<br>Arlington, VA 22217 | | 12. REPORT DATE<br>April 1982 |
| | | 13. NUMBER OF PAGES<br>69 |
| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) | | 15. SECURITY CLASS. (of this report) |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |
| 16. DISTRIBUTION STATEMENT (of this Report)<br>Approved for public release: distribution unlimited | | |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) | | |
| 18. SUPPLEMENTARY NOTES | | |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number) | | |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number)<br>In this first of two papers, computable a posteriori estimates of the space discretization error in the finite element method of lines solution of parabolic equations are analyzed for time-independent space meshes. The effectivity of the error estimator is related to conditions on the solution regularity, mesh family type, and asymptotic range for the mesh size. For clarity the results are limited to a model problem in which piecewise linear elements in one space dimension are used. The results extend straightforwardly to systems of equations and higher order elements in one space dimen- | | |

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

28. sion, while the higher dimensional case requires additional considerations. The theory presented here provides the basis for the analysis and adaptive construction of time-dependent space meshes, which is the subject of the second paper. Computational results show that the approach is practically very effective and suggest that it can be used for solving more general problems.

# THE FINITE ELEMENT METHOD FOR PARABOLIC EQUATIONS,

## I. A POSTERIORI ERROR ESTIMATION

Technical Note BN-983

by

M. Bieterman [1]

and

I. Babuška [2]

1 Laboratory of Applied Studies
Division of Computer Research and Technology
National Institutes of Health
Bethesda, MD 20205

2 Institute for Physical Science and Technology
and Department of Mathematics
University of Maryland
College Park, MD 20742

The work of the second author was partially supported by ONR Contract N00014-77-C-0623.

## Abstract

In this first of two papers, computable a posteriori estimates of the space discretization error in the finite element method of lines solution of parabolic equations are analyzed for time-independent space meshes. The effectivity of the error estimator is related to conditions on the solution regularity, mesh family type, and asymptotic range for the mesh size. For clarity the results are limited to a model problem in which piecewise linear elements in one space dimension are used. The results extend straightforwardly to systems of equations and higher order elements in one space dimension, while the higher dimensional case requires additional considerations. The theory presented here provides the basis for the analysis and adaptive construction of time-dependent space meshes, which is the subject of the second paper. Computational results show that the approach is practically very effective and suggest that it can be used for solving more general problems.

## 1. Introduction

In recent years interest has grown in the method of lines (MOL) approach for the numerical solution of time-dependent partial differential equations (PDEs) arising in biology, rheology, structural and fluid mechanics, and many other fields. In this approach a problem is first discretized in space, for example, by a Galerkin or difference method. This results in a system of ordinary differential equations (ODEs) which can be efficiently solved by one of the available variable-step, variable-order ODE packages. This segmented approach is obviously practical, and modern ODE integrators are generally reliable in estimating and controlling the error due to the time discretization. However, one is usually interested in the total error, i.e. the error of the approximate solution with respect to the exact solution of the original PDE. The advantages of highly accurate ODE integrators are diminished if the error due to the space discretization is large. Those available programs implementing both segments of the MOL approach (cf. survey in [16]) generally have no facility to estimate and control both components of the error.

Recently a MOL procedure, which we shall call the finite element method of lines (FEMOL), was proposed for the solution of parabolic PDEs (cf. [1]). Piecewise linear finite elements in one space dimension were employed in a model problem, which consisted of a second order homogeneous PDE, with smooth and compatible initial data. Using the theory and practice of a posteriori error estimates developed for elliptic PDEs in [2], [3], [4], and [5], an estimator for the space discretization error was constructed. This estimator and a novel adaptive time discretization scheme were utilized to estimate the total error in the FEMOL in the case of time-independent space meshes.

We shall restrict our attention here to the space discretization error in the FEMOL, i.e. the total error, assuming that the resulting ODE system can be solved exactly. In any given application this of course cannot be accomplished, and the effects of the errors in the ODEs must be taken into account. The purpose of this paper is to completly isolate the space component of the error and means of estimating it. By keeping the analysis independent of the time discretization scheme, none of the state of the art ODE solvers are excluded a priori when dealing with the important question of how to properly balance the two components of the error.

In this first of two papers, it is assumed that the space mesh remains fixed throughout the time evolution of the problem. Using piecewise linear elements in one space dimension, we consider the model problem of a second order nonhomogeneous parabolic PDE, with initial data not necessarily smooth or compatible. For related results with higher order elements in the setting of coupled systems of equations in one space dimension see [8].

The results given here are extended in a second paper [7] to analyze means of estimating the space discretization error when the space mesh is allowed to change discontinuously in time. Based upon these estimates, a procedure is given there for the control of the space discretization error by the adaptive construction of space meshes during the solution process. This algorithm is then tested in some computational examples.

The effectivity of the space discretization error estimator constructed in [1] depends on properties of the exact solution in the space-time domain and is asymptotic in nature. We give sufficient conditions and a priori bounds on the asymptotic range for this estimator to work. These bounds are overly pessimistic for any given problem, and we shall see in some examples that the results are much better than the theory predicts. Under less restrictive conditions and for all mesh partitions, it is also shown that a modification of our principal error estimator can be expected to perform well. This modification takes more fully into account the case where an exact solution has a strong time-dependence with respect to the size of the mesh used.

This paper consists of the following sections. In section 2 notation, mathematical preliminaries, and the model problem are introduced. In section 3 certain regularity classes are defined and an accompanying theorem is proven, showing that the model problem is well-posed. Based upon this proof and result, in section 4 a priori estimates needed in section 5 are given. Our a posteriori error estimates are presented in sections 5 and 6. Section 7 consists of the description of computational procedures used and examples.

Many of the required function spaces, results, and techniques used in sections 2-6 to prove our results are somewhat technical. In order to facilitate a first reading of the paper, we present here some basic notions and a sample of our computational results. By then skipping directly to section 7, the reader should have a good overview of the main ideas presented.

Let $u = u(t,x)$ be the solution of the model problem

$$(1.1)\quad \begin{cases} u_t = f - Lu = f(t,x) + [a(x)u_x]_x - b(x)u; & 0 < x < 1, \quad t > 0, \\ u(t,x) = g(t,x); & x = 0,1, \quad t > 0, \\ u(0,x) = u_0(x); & 0 < x < 1, \end{cases}$$

where $a > 0$, $b \geq 0$, $u_0$, $g$, and $f$ are given functions.

The FEMOL solution of eqs. (1.1) is accomplished by partitioning the interval $(0,1)$ according to $\Delta^{(N)} = \{0 = x_0 < x_1 < \cdots < x_N = 1\}$ and determining the unique function $U = U(t,x)$, which is piecewise linear in $x$ and smooth in $t$, and satisfies a weak, or integral form of eqs. (1.1). For each $t$, $U$ is characterized by the vector of its values $\{U(t,x_j)\}_{j=1,N-1}$, which is the solution of an $N-1$ dimensional ODE system, assumed here to be exactly solvable. In the computational examples, the ODE systems were solved by using successively smaller time discretization error tolerances, the process being terminated when all relevant quantities experienced only insignificant changes in the time intervals of interest.

For each $t > 0$ the goal is to estimate

$$(1.2) \qquad |||e(t,\cdot)||| = \left\{\int_0^1 a(x)e_x^2(t,x)dx\right\}^{1/2},$$

where $e = u - U$ is the space discretization error.

In order to estimate $|||e(t,\cdot)|||$, we define the local error indicators $\{\eta_j(t)\}_{j=1,N}$ by

$$(1.3) \qquad \eta_j^2(t) = \frac{|x_j-x_{j-1}|^2}{12\, a(\frac{x_{j-1}+x_j}{2})}\int_{x_{j-1}}^{x_j} r^2(t,x)dx; \qquad 1 \le j \le N,$$

where the residual $r(t,x) \equiv (U_t+LU-f)(t,x)$ is well-defined and computable for each $t > 0$ on each subinterval $(x_{j-1},x_j)$, given $U$.

The principal a posteriori error estimator $E(\cdot)$ is then defined by

$$(1.4) \qquad E(t) = \left\{\sum_{j=1}^{N} \eta_j^2(t)\right\}^{1/2}; \qquad t > 0.$$

To assess the performance of $E(\cdot)$, we define the effectivity ratio

$$(1.5) \qquad \Theta(t) = E(t)/|||e(t,\cdot)|||; \qquad t > 0.$$

$\mathcal{E}(\cdot)$ is an effective estimator if there exists a reasonable constant $C$ (depending on the admissible solution class for eqs. (1.1) and the admissible class of mesh partitions to which $\Delta^{(N)}$ belongs, but not on the magnitude of the data $u_0$, $f$, or $g$) such that

$$(1.6) \qquad 1/\Theta(t) \leq C; \qquad \forall t > 0.$$

Moreover, it is desirable that

(1.7) $\Theta(t) \to 1$ as the partitions of $(0,1)$ are refined.

In one of the examples of section 7, the exact solution of eqs. (1.1) is taken to be

$$(1.8) \qquad u(t,x) = \frac{1}{2} + \frac{1}{2}\tanh\,[2\beta(x-10t)]; \qquad \beta = 20,$$

which is a wave with approximate front width $\beta^{-1}$ that moves in the positive $x$-direction at speed 10 (cf. figure 1(a), (b),(c)). For each of various values of $N$, the FEMOL was implemented in this example by defining the partition $\Delta^{(N)}$ according to $x_j = j/N$; $j = 0,N$. The tables below illustrate for various $t$ the performance of $\mathcal{E}(t)$ as compared with the relative error $E_{REL}(t)$, defined by

$$(1.9) \qquad E_{REL}(t) = |||e(t,\cdot)|||/|||u(t,\cdot)|||.$$

TABLE 1. $E_{REL}(t)$

| t \ N | 20 | 40 | 80 | 160 |
|---|---|---|---|---|
| .032 | .4325 | .2634 | .1308 | .0649 |
| .056 | .5697 | .2431 | .1309 | .0649 |
| .060 | .6437 | .2818 | .1309 | .0649 |
| .088 | .3473 | .2673 | .1309 | .0649 |

TABLE 2. $\Theta(t)$

| t \ N | 20 | 40 | 80 | 160 |
|---|---|---|---|---|
| .032 | 1.320 | .995 | .989 | .996 |
| .056 | .776 | 1.076 | .989 | .996 |
| .060 | .806 | .903 | .988 | .996 |
| .088 | 1.685 | .961 | .987 | .996 |

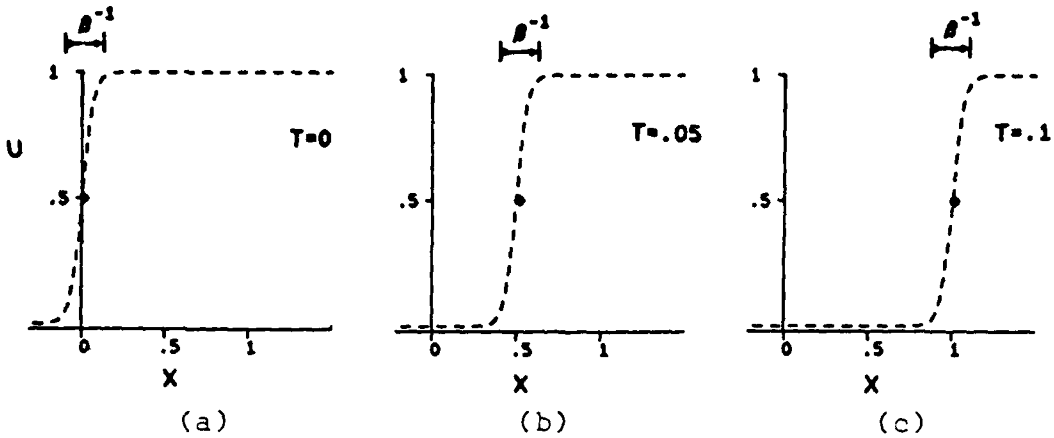


Figure 1. Solution u(T,X) in Eq. (1.8) vs. X for various T.

The last three columns of Table 1 illustrate for each $t$ the expected linear rate of convergence for $|||e(t,\cdot)|||$. For $N = 20$ the mesh used is not fine enough to resolve the wavefront and the errors oscillate about the 50% mark.

The numbers in Table 2 clearly illustrate the effectivity of estimating $|||e(t,\cdot)|||$ with $\mathcal{E}(t)$. While values for the constant C in (1.6) on the order of 10 or more would be acceptable in many applications, for none of the entries in Table 2 is such a pessimistic bound realized. We see that for about 13% or less error, the estimator $\mathcal{E}(\cdot)$ differs from $|||e(\cdot,\cdot)|||$ by no more than about 1%, a truly remarkable result.

These and other computational results are discussed more fully in section 7. We again mention that the estimator $\mathcal{E}(\cdot)$ can be used in an algorithm, where decisions concerning modifications of the space mesh are made adaptively during the solution process. Provided that proper constraints are imposed, this then leads to an implementation of the FEMOL which is robust enough to adaptively modify meshes in the case of localized activity, such as for the travelling wave above, and yet is still efficient enough to handle problems where little or no mesh modification is necessary.

## 2. Mathematical Preliminaries and the Model Problem

Throughout this paper we shall use the convention that the variables $i$, $j$, $\ell$, $m$, $N$, $p$, and $q$ take only integer values.

Let $J$ be a bounded open interval in $\mathbb{R}^1$, $\bar{J}$ its closure, and $V$ be an arbitrary Hilbert space. We denote by $C^0(\bar{J};V)$ those $V$-valued functions which are continuous on $\bar{J}$, $C^\infty(\bar{J};V)$ the subset of infinitely differentiable functions on $J$ for which all derivatives have continuous extensions on $\bar{J}$, and $C_0^\infty(J;V)$ those functions in $C^\infty(\bar{J};V)$ with compact support in $J$. When $V = \mathbb{R}^1$ we suppress $V$ and write $C^0(\bar{J})$, $C^\infty(\bar{J})$, $C_0^\infty(J)$.

Let $I = (0,1)$ be the unit open interval in $\mathbb{R}^1$, $\bar{I}$ its closure, and $[0,T)$ be a bounded half open interval in $\mathbb{R}^1$. For $\phi,\psi \in C^\infty(\bar{I})$, set

$$<\phi,\psi> \;=\; \int_I \phi(x)\psi(x)dx \quad\text{and}\quad \|\phi\|_0^2 \;=\; <\phi,\phi>.$$

For each $i \geq 0$ the Sobolev space $H^i\{H_0^i\}$ is the completion of $C^\infty(\bar{I})\{C_0^\infty(I)\}$ with respect to the norm

$$\|\phi\|_i^2 \;=\; \sum_{p=0}^{i} \left\|\frac{d^p\phi}{dx^p}\right\|_0^2 ,$$

and for each real $s \geq 0$, $H^s$ with norm $\|\cdot\|_s$ is a Hilbert space which is defined by interpolation as in [15].

Let $a$ and $b$ be sufficiently smooth functions on $\bar{I}$ for which

$$
\left.\begin{array}{ll}
(2.1) & 0 < \underline{a} \le a(x) \le \bar{a} < \infty \\
(2.2) & 0 \le b(x) \le \bar{b} < \infty
\end{array}\right\} \qquad \forall x \in \bar{I}.
$$

For $u,v \in H_0^1$ we define the bilinear form

$$(2.3) \qquad B(u,v) = \langle au_x, v_x \rangle + \langle bu, v \rangle.$$

By (2.1) and (2.2) $\exists$ positive constants $C_1, C_2$ such that

$$(2.4) \qquad C_2 \|w\|_1^2 \le B(w,w) \le C_1 \|w\|_1^2 \qquad \forall w \in H_0^1,$$

and $\|w\|_E^2 \equiv B(w,w)$ defines a norm equivalent to $\|\cdot\|_1$ on $H_0^1$.

The operator $L$ defined by

$$Lu = -(au_x)_x + bu$$

with domain $\mathcal{D}(L) = H^2 \cap H_0^1$ is the unique positive self-adjoint operator induced by $B$. Using spectral representations (cf. [9], [13], [14]) we can define fractional powers of $L$. Results in [11] show that for $\alpha \ge 0$, $\alpha/2 \ne$ integer $+\frac{1}{4}$

$$(2.5) \qquad \mathcal{D}(L^{\alpha/2}) = \{v \in H^\alpha : L^p v(x)|_{x=0,1} = 0;$$
$$\forall p \text{ satisfying } 0 \le p < \alpha/2 - 1/4\}$$

and, as in (2.4),

(2.6) the spectral norm $\|L^{\alpha/2}\cdot\|_0$ is equivalent to $\|\cdot\|_\alpha$ on $\mathcal{D}(L^{\alpha/2})$.

Let $P$ be a family of mesh partitions of $I$ and

$$\Delta = \{0 = x_0 < x_1 < \cdots < x_{N(\Delta)} = 1\} \in P.$$

For $j = 1, N(\Delta)$ we write

$$I_j = (x_{j-1}, x_j), \qquad h_j = x_j - x_{j-1}, \quad \text{and}$$

$$\underline{h}(\Delta) = \min_{1 \le j \le N(\Delta)} h_j\,, \qquad h(\Delta) = \max_{1 \le j \le N(\Delta)} h_j\,.$$

$P$ is said to be a $\kappa(\lambda)$-regular family if $\exists$ constants $\kappa \geq 1$, $\lambda > 0$ such that

$$\underline{h}(\Delta) \geq \lambda h^{\kappa}(\Delta) \qquad \forall \Delta \in P. \tag{2.7}$$

For $v \in H^0$, $w \in H_0^1$, $j = 1, N(\Delta)$, and $\psi \in C^0(\bar{I}_j)$ we shall write

$$\|v\|_{0,I_j}^2 = \int_{I_j} v^2(x)dx,$$

$$\|w\|_{E,I_j}^2 = \int_{I_j} [a(x)w_x^2(x) + b(x)w^2(x)]dx, \quad \text{and}$$

$$\psi_j = \psi\left(\frac{x_{j-1}+x_j}{2}\right).$$

For each $\Delta \in P$, $S(\Delta) \subset H_0^1$ is the subset of functions whose restrictions to any $I_j$ are linear. The operators $P_0^{\Delta}$, $P_1^{\Delta}$, $L_{\Delta}$ mapping $H^0$, $H_0^1$, $S(\Delta)$, respectively, to $S(\Delta)$ are defined by

$$(2.8)\qquad \begin{cases} \langle P_0^\Delta v - v, \phi\rangle = 0 \\ B(P_1^\Delta v - v, \phi) = 0 \\ \langle L_\Delta v, \phi\rangle = B(v,\phi) \end{cases} \qquad \forall \phi \in S(\Delta).$$

$L_\Delta$ is a positive self-adjoint operator on $S(\Delta)$ satisfying

$$(2.9)\qquad L_\Delta(P_1^\Delta u) = P_0^\Delta(Lu) \qquad \forall u \in \mathcal{D}(L).$$

By (2.1) and (2.2) $\exists$ a positive constant $C$ such that $\forall \phi \in S(\Delta)$

$$(2.10)\qquad \|L_\Delta^{-1}\phi\|_0 \le C\|L_\Delta^{-1}\phi\|_E \le C^2\|\phi\|_0,$$

and if $\mathcal{P}$ is $\kappa(\lambda)$-regular

$$(2.11)\qquad \|L_\Delta \phi\|_0 \le (C/\lambda)h^{-\kappa}(\Delta)\|\phi\|_E.$$

The following lemma is well-known (cf. [6, Chap. 4]).

Lemma 2.1. There exists a positive constant $C$ such that for each $\Delta \in \mathcal{P}$

$$(2.12)\qquad \|(I-P_1^\Delta)v\|_s \le Ch^{\sigma-s}(\Delta)\|v\|_\sigma; \qquad \forall v \in \mathcal{D}(L^{\sigma/2}),\ 0 \le s \le 1 \le \sigma \le 2,$$

$$(2.13)\qquad \|(I-P_0^\Delta)v\|_0 \le Ch^{\sigma}(\Delta)\|v\|_\sigma; \qquad \forall v \in \mathcal{D}(L^{\sigma/2}),\ 0 \le \sigma \le 2,\ \sigma \ne 1/2.$$

We shall consider the following model problem.

$$(2.14)\quad \begin{cases} \frac{\partial}{\partial t} u(t,x) + Lu(t,x) = f(t,x); & t,x \in (0,T) \times I, \\ u(t,0) = 0 = u(t,1); & t \in (0,T), \\ u(0,x) = u_0(x); & x \in I. \end{cases}$$

The data $u_0$ and $f$ are such that the solution $u(;u_0,f)$ of eqs. (2.14) resides in a certain regularity class $\mathcal{H}$, which will be specified in section 3.

The weak form of eqs. (2.14) is: find $u:(0,T) \to H_0^1$ satisfying

$$(2.15)\quad \begin{cases} \langle u_t(t),\phi\rangle + B(u(t),\phi) = \langle f(t),\phi\rangle; & \forall t,\phi \in (0,T) \times H_0^1, \\ u(0) = u_0. \end{cases}$$

For each $\Delta \in \mathcal{P}$, the FEMOL solution $U^\Delta : [0,T) \to S(\Delta)$ is defined by

$$(2.16)\quad \begin{cases} \langle U_t^\Delta(t),\phi\rangle + B(U^\Delta(t),\phi) = \langle f(t),\phi\rangle; & \forall t,\phi \in (0,T) \times S(\Delta), \\ U^\Delta(0) = U_0^\Delta \in S(\Delta), \end{cases}$$

where $U_0^\Delta$ is chosen to approximate $u_0$ and is specified in section 4. Eqs. (2.16) may be rewritten as

$$(2.17)\quad \begin{cases} U_t^\Delta(t) + L_\Delta U^\Delta(t) = P_0^\Delta f(t); & t \in (0,T), \\ U^\Delta(0) = U_0^\Delta, \end{cases}$$

and for reasonable data $u_0$, $f$ are equivalent to a uniquely solvable $N(\Delta) - 1$ dimensional ODE initial value problem.

The following discrete smoothing property can easily be verified (cf. [9]).

Lemma 2.2. Let $U^\Delta = U^\Delta(;U_0^\Delta,0)$ denote the solution of eqs. (2.17) with $P_0^\Delta f \equiv 0$. Then $\exists$ a positive constant $C$ such that for $i = 0$ or $1$

$$\text{(2.18)}\qquad \left\|\frac{\partial^j}{\partial t^j} U^\Delta(t)\right\|_i \le Ct^{-(i+2j)/2}\,\|U_0^\Delta\|_0; \qquad \forall t > 0,\quad j \ge 0.$$

Remark. When there is no threat of confusion, we shall often drop $\Delta$ when writing $S$, $h$, $P_0$, $P_1$, $U$, and $U_0$.

Our work requires the use of the nonisotropic Sobolev spaces $H^{2s,s}$, which we now define. For nonnegative $\ell$ and $m$, let $H^\ell(J;H^m)$ be the completion of $C^\infty(\bar{J};H^m)$ with respect to the norm

$$\left(\|u\|^J_{H^\ell(H^m)}\right)^2 = \sum_{j=0}^{\ell} \int_J \left\|\frac{\partial^j}{\partial t^j} u(t,\cdot)\right\|_m^2 dt,$$

and define $H^{2\ell,\ell}(J) = H^0(J;H^{2\ell}) \cap H^\ell(J;H^0)$ with norm

$$\left(\|u\|^J_{2\ell,\ell}\right)^2 = \left(\|u\|^J_{H^0(H^{2\ell})}\right)^2 + \left(\|u\|^J_{H^\ell(H^0)}\right)^2 .$$

For each real $s \ge 0$ the space $H^{2s,s}(J)$ with norm $\|\cdot\|^J_{2s,s}$ is a Hilbert space which is defined by interpolation as in [15]. We state in convenient forms some well-known trace and embedding results.

Theorem 2.1. (cf. [15, Thms. 1.2.3, 1.4.1, Prop. 4.2.3])

Let $s \geq 0$, $v \in H^{2s,s}(J)$, and $j$ satisfy $0 \leq j \leq s$. Then $\frac{\partial^j}{\partial t^j} v \in H^{2(s-j),(s-j)}(J)$ and $\exists$ a positive constant $C$ such that

$$\left\| \frac{\partial^j}{\partial t^j} v \right\|^J_{2(s-j),(s-j)} \leq C \|v\|^J_{2s,s} .$$

Theorem 2.2. (cf. [15, Thms. 1.3.1, 1.4.2])

Let $s > 1/2$, $v \in H^{2s,s}(J)$, and $j \in [0, s-1/2)$. Then $\frac{\partial^j}{\partial t^j} v \in C^0(\bar{J}; H^{2(s-j-1/2)})$ and $\exists$ a positive constant $C$ such that

$$\sup_{t \in \bar{J}} \left\| \frac{\partial^j}{\partial t^j} v(t,\cdot) \right\|_{2(s-j-1/2)} \leq C \|v\|^J_{2s,s} .$$

Theorem 2.3. (cf. [15, Thm. 4.2.1, Prop. 4.2.2])

Let $s > 3/4$, $v \in H^{2s,s}(J)$, $t^* \in \bar{J}$, and $j$ and $p$ satisfy $0 \leq j + p/2 < s - 3/4$. Then

$$\frac{\partial^j}{\partial t^j}\left( \frac{\partial^p}{\partial x^p} v(t,x) \Big|_{x=0,1} \right)\Big|_{t=t^*} = \frac{\partial^p}{\partial x^p}\left( \frac{\partial^j}{\partial t^j} v(t,x) \Big|_{t=t^*} \right)\Big|_{x=0,1} .$$

We conclude this section by noting a useful partial extension of Lemma 2.1.

Lemma 2.3. There exists a positive constant $C$ such that for each $\Delta \in \mathcal{P}$

$$\|(I-P_1^\Delta)\frac{\partial^j}{\partial t^j} v\|^J_{0,0} \leq C h^\sigma(\Delta) \|v\|^J_{2\mu,\mu}; \quad \forall v \in H^{2\mu,\mu}(J) \cap H^0(J; H_0^1),$$

$$\mu = j + \sigma/2, \quad j \geq 0, \quad 1 \leq \sigma \leq 2.$$

Proof. Set $\mathcal{V} = H^0(J;\mathcal{D}(L^{\sigma/2}))$. Since $C^\infty(\bar{J};\mathcal{D}(L^{\sigma/2}))$ is dense in the Hilbert space $\mathcal{V}$, $I - P_1^\Delta$ can be continuously extended to an operator mapping $\mathcal{V}$ to $H^{0,0}(J)$. By (2.12) we have

$$(2.19) \qquad \|(I-P_1^\Delta)w\|_{0,0}^J \le Ch^\sigma(\Delta)\|w\|_{H^0(\mathcal{D}(L^{\sigma/2}))}^J; \qquad \forall w \in \mathcal{V}.$$

It follows from the assumptions of the lemma, (2.5), and Thms. 2.1-2.3 that $w = \frac{\partial^j}{\partial t^j}v \in \mathcal{V}$.

The application of (2.6) and Thm. 2.1 in (2.19) completes the proof.

## 3. Parabolic Regularity

Throughout the remainder of this paper, $\alpha$ and $\beta$ denote arbitrary but fixed real numbers related to the data in eqs. (2.14) and satisfy

$$(3.1) \quad \alpha > 1, \quad -1 \le \beta \le \alpha, \quad \beta, \alpha, \alpha/2 \ne \text{integer} + 1/2.$$

For real $\sigma$, $[\sigma]$ denotes the integral part of $\sigma$ and $\sum'^{q}_{j=p} \equiv 0$ for $q < p$.

We begin by defining for $v \in H^{\beta+1}$, $g \in H^{\alpha,\alpha/2}(0,T)$ the $\beta$-order compatibility measure

$$(3.2) \quad M^{CR}_{\beta}(v,g) = \sum_{x=0,1} \sum'^{[\beta/2+1/4]}_{\ell=0} \left| L^{\ell} v(x) + \sum'^{\ell}_{p=1} (-1)^p L^{\ell-p} \frac{\partial^{p-1}}{\partial t^{p-1}} g(0,x) \right|,$$

which by Thms. 2.1 - 2.3 is well-defined.

The following facts concerning solutions $u(;\varphi,f)$ of eqs. (2.14) with initial data $\varphi$ and right hand side $f$ are well-known.

Theorem 3.1. (Existence; cf. [15, Thm. 3.4.1]) Let $\varphi \in H^0$, $f \in H^{0,0}(0,T)$. Then eqs. (2.14) have a unique solution $u(;\varphi,f) \in H^0((0,T);H^1_0)$.

Theorem 3.2. (Regularity; cf. [15, Thms. 4.53, 4.62]) Let $\varphi \in H^{\alpha+1}$, $f \in H^{\alpha,\alpha/2}(0,T)$, and $M^{CR}_{\alpha}(\varphi,f) = 0$. Then $u(;\varphi,f) \in H^{\alpha+2,(\alpha+2)/2}(0,T)$ and $\exists$ a positive constant $C$ such that

$$\|u\|^{(0,T)}_{\alpha+2,(\alpha+2)/2} \le C\{\|\varphi\|_{\alpha+1} + \|f\|^{(0,T)}_{\alpha,\alpha/2}\}.$$

Lemma 3.1. (cf. [13, p. 489], [14], and (2 .5))

Let $f \equiv 0$ and $\varphi \in \mathcal{D}(L^{(\beta+1)/2})$. Then $\forall t > 0$ $u(t;\varphi,0) \in \mathcal{D}(L^{(\alpha+1)/2})$ and $\exists$ a positive constant $C$ such that

$$\|u(t;\varphi,0)\|_{\alpha+1} \le C\ t^{-(\alpha+1)/2}\|\varphi\|_{\beta+1}.$$

We therefore have

Theorem 3.3. (Parabolic Smoothing)

Let $\varphi$, $u(;\varphi,0)$ be as in Lemma 3.1 and $0 < T_* < T$. Then

(3.3) $u(;\varphi,0) \in H^{\alpha+2,(\alpha+2)/2}(T_*,T)$ and $\exists$ a positive constant $C$ such that

$$\|u(;\varphi,0)\|^{(T_*,T)}_{\alpha+2,(\alpha+2)/2} \le C\ T_*^{-(\alpha+1)/2}\|\varphi\|_{\beta+1}. \tag{3.4}$$

Proof. By Lemma 3.1 $u(T_*;\varphi,0) \in \mathcal{D}(L^{(\alpha+1)/2})$, which is equivalent to $M^{CR}_{\alpha}(u(T_*;\varphi,0),0) = 0$. By applying Thm. 3.2 to the interval $(T_*,T)$ we have (3.3) and

$$\|u(;u(T_*;\varphi,0),0)\|^{(T_*,T)}_{\alpha+2,(\alpha+2)/2} \le C\|u(T_*;\varphi,0)\|_{\alpha+1}$$

$$\text{by Lemma 3.1} \quad \le C\ T_*^{-(\alpha+1)/2}\|\varphi\|_{\beta+1}.$$

We now define certain regularity classes for data and solutions of eqs. (2.14) in order to combine Thms. 3.2, 3.3 for use in Section 4. Set

$$(3.5)\quad \mathbb{D}^{\alpha,\beta}(0,T) = \{v,g \in H^{\beta+1} \times H^{\alpha,\alpha/2}(0,T) : M_{\beta}^{CR}(v,g) = 0\}.$$

For $v,g \in \mathbb{D}^{\alpha,\alpha}(0,T)$ and $0 \le T_* < T$ set

$$(3.6)\quad \overline{C}_{\alpha}(\alpha,T_*,T,v,g) = \|v\|_{\alpha+1} + \|g\|_{\alpha,\alpha/2}^{(0,T)},$$

and for $v,g \in \mathbb{D}^{\alpha,\beta}(0,T)$ with $\beta < \alpha$ and $0 < T_* < T$ set

$$(3.7)\quad \overline{C}_{\alpha}(\beta,T_*,T,v,g) = (1 + T_*^{-(\alpha+1)/2})(M_{\alpha}^{CR}(0,g) + \|v\|_{\beta+1}) + \|g\|_{\alpha,\alpha/2}^{(0,T)}.$$

For $0 \le T_* < T$ define

$$(3.8)\quad H^{\alpha,\beta}(T_*,T) = \{u \in H^{\alpha+2,(\alpha+2)/2}(T_*,T) : u = u(;v,g) \text{ solves (2.14) for some } v,g \in \mathbb{D}^{\alpha,\beta}(0,T)\},$$

and for $u(;v,g) \in H^{\alpha,\beta}(T_*,T)$ set

$$(3.9)\quad C_{\alpha}(\beta,T_*,T,u) = \overline{C}_{\alpha}(\beta,T_*,T,v,g).$$

By properly decomposing given initial data we shall show

<u>Theorem 3.4</u>. Let $0 \le T_* < T$ and $u_0,f \in \mathbb{D}^{\alpha,\beta}(0,T)$, where $\beta = \alpha$ if $T_* = 0$. Then eqs. (2.14) have a unique solution $u(;u_0,f) \in H^{\alpha,\beta}(T_*,T)$ and $\exists$ a positive constant $C$ such that

$$(3.10)\quad \|u\|^{(T_*,T)}_{\alpha+2,(\alpha+2)/2} \le C\,\overline{C}_\alpha(\beta,T_*,T,u_0,f) = C\,C_\alpha(\beta,T_*,T,u).$$

Remark. The decomposition used in the proof of Thm. 3.4 will be used to demonstrate the a priori estimates of Section 4.

Proof. The outline of the proof is as follows. Through a sequence of elliptic boundary value problems we construct functions $u_0^1$ and $u_0^2$ with the following properties:

$$(3.11)\qquad u_0^1 + u_0^2 = u_0,$$

$$(3.12)\qquad u_0^1 \in H^{\alpha+1} \quad \text{and} \quad M_\alpha^{CR}(u_0^1,f) = 0,$$

$$(3.13)\qquad u_0^2 \in \mathcal{D}(L^{(\beta+1)/2}),$$

and for which $\exists$ positive constants $C_1$, $C_2$ such that

$$(3.14)\quad \|u_0^1\|_{\alpha+1} \le C_1 \begin{cases} \|u_0\|_{\alpha+1}, & \text{if } \beta = \alpha, \\ M_\alpha^{CR}(0,f), & \text{if } \beta < \alpha, \end{cases}$$

$$(3.15)\quad \|u_0^2\|_{\beta+1} \le C_2 \begin{cases} 0, & \text{if } \beta = \alpha, \\ M_\alpha^{CR}(0,f) + \|u_0\|_{\beta+1}, & \text{if } \beta < \alpha. \end{cases}$$

By linearity the solution $u(;u_0,f)$ (whose existence is given by Thm. 3.1) can be decomposed as

$$(3.16)\quad u = u^1 + u^2; \quad u^1 = u(;u_0^1,f), \quad u^2 = u(;u_0^2,0).$$

It then follows by Thms. 3.2, 3.3, (3.6)-(3.9), (3.16), and the triangle inequality that $u \in H^{\alpha,\beta}(T_*,T)$ and

$$(3.17)\quad \|u\|_{\alpha+2,(\alpha+2)/2}^{(T_*,T)} \le \|u^1\|_{\alpha+2,(\alpha+2)/2}^{(0,T)} + \|u^2\|_{\alpha+2,(\alpha+2)/2}^{(T_*,T)}$$

$$\le C\{\|u_0^1\|_{\alpha+1} + \|f\|_{\alpha,\alpha/2}^{(0,T)} + T_*^{-(\alpha+1)/2} \|u_0^2\|_{\beta+1}\}$$

by (3.14) and (3.15) $\quad \le C\begin{cases} \|u_0\|_{\alpha+1} + \|f\|_{\alpha,\alpha/2}^{(0,T)}; & \text{if } \beta = \alpha, \\ (1+T_*^{-(\alpha+1)/2})M_\alpha^{CR}(0,f) + T_*^{-(\alpha+1)/2} \|u_0\|_{\beta+1} + \|f\|_{\alpha,\alpha/2}^{(0,T)}; & \text{if } \beta < \alpha, \end{cases}$

by (3.6)-(3.9) $\quad \le C_\alpha(\beta,T_*,T,u).$

It remains therefore to construct $u_0^1$ and $u_0^2$ satisfying (3.11)-(3.15).

$$(3.18)\quad \begin{cases} \text{If } \beta = \alpha, \text{ set } u_0^1 = u_0 \text{ and } u_0^2 = 0, \text{ and} \\ \text{if } \beta < \alpha, \text{ set } u_0^1 = \psi_\alpha \text{ and } u_0^2 = u_0 - \psi_\alpha, \end{cases}$$

where with $i = [\alpha/2 + 1/4] \ge 0$, $\psi_\alpha \equiv z_i$, and the sequence $\{z_j\}_{j=-1,i}$ is defined by:

$z_{-1} \equiv 0$ and for $j = 0,i$ $z_j$ solves

$$(3.19)\quad \begin{cases} Lz_j(x) = z_{j-1}(x) & ;x \in I, \\ z_j(x) = -\sum_{p=1}^{i-j}{}'(-1)^p L^{i-j-p} \frac{\partial^{p-1}}{\partial t^{p-1}} f(0,x); & x = 0,1. \end{cases}$$

By Thms. 2.1-2.3 and elliptic regularity theory we have that $\psi_\alpha \equiv z_i \in H^{\alpha+1} \cap \mathcal{D}(L)$ and $\exists$ a positive constant $C$ such that

$$(3.20)\quad \|\psi_\alpha\|_{\alpha+1} \leq C \sum_{x=0,1} \sum_{j=0}^{i-1}{}' |z_j(x)| = C\, M_\alpha^{CR}(0,f).$$

Properties (3.11)-(3.18) are easily verified, thus completing the proof of Thm. 3.4.

We now define some subclasses of $H^{\alpha,\beta}(T_*,T)$ which are needed in Section 5.

For $\delta_1 > 0$ set

$$(3.21)\quad H^{\alpha,\beta}_{\delta_1}(T_*,T) = \{u \in H^{\alpha,\beta}(T_*,T) : u \text{ satisfies } I(\delta_1)\},$$

and for $\delta_1 > 0$, $\delta_2 > 0$, $\alpha > 3$ set

$$(3.22)\quad H^{\alpha,\beta}_{\delta_1,\delta_2}(T_*,T) = \{u \in H^{\alpha,\beta}_{\delta_1}(T_*,T) : u \text{ satisfies } II(\delta_2)\},$$

where the properties $I(\delta_1)$ and $II(\delta_2)$ are defined as follows.

We say that $u \in H^{\alpha,\beta}(T_*,T)$ satisfies $I(\delta_1)$ if

(3.23) for each $t \in [T_*,T]$, $\|u_{xx}(t,\cdot)\|_0 > 0$ and

$$\sup_{t\in[T_*,T]} \frac{C_\alpha(\beta,T_*,T,u)}{\|u_{xx}(t,\ )\|_0} \le \delta_1 .$$

We say that $u \in H^{\alpha,\beta}(T_*,T)$ satisfies $II(\delta_2)$ if

(3.24) for each $t \in [T_*,T]$ $\exists$ an open interval $I_t \subset I$ such that

(3.25) $g_L(t,u) \equiv \inf_{x\in I_t} \{|u_{xx}(t,x)| \cdot |I_t|^{1/2}\} > 0$ and

(3.26) $$\sup_{t\in[T_*,T]} \left\{\frac{g_U(t,u)}{g_L(t,u)}\right\} \le \delta_2, \text{ where}$$

(3.27) $g_U(t,u) \equiv \sup_{x\in I}\{|u_{xx}(t,x)| + |u_{xxx}(t,x)|\}$.

Remarks. From Thms. 2.2, 3.4, and the Sobolev Embedding Theorem we have that (3.23)-(3.27) are well-defined, since for $\alpha > 1$, $u \in H^{\alpha,\beta}(T_*,T) \Rightarrow u \in C^0([T_*,T];H^2)$, and for $\alpha > 3$, $u \in H^{\alpha,\beta}(T_*,T) \Rightarrow u_{xxx} \in C^0([T_*,T] \times \bar{I})$. Membership in these subclasses does not depend on absolute magnitude.

For $u \in H^{\alpha,\beta}_{\delta_1}\{H^{\alpha,\beta}_{\delta_1,\delta_2}\}$, it follows that $C \cdot u \in H^{\alpha,\beta}_{\delta_1}\{H^{\alpha,\beta}_{\delta_1,\delta_2}\}$ for any nonzero constant $C$ and the same $\delta_1\{\delta_1$ and $\delta_2\}$.

## 4. A Priori Estimates

Let $u(t) = u(t;u_0,f)$ and $U(t) = U^{\Delta}(t;U_0^{\Delta},P_0^{\Delta}f)$ denote the solutions of eqs. (2.14) and (2.17), respectively. Define

$$(4.1)\qquad \begin{cases} e(t) = u(t) - U(t), \\ \rho(t) = P_1u(t) - u(t), \\ \theta(t) = P_1u(t) - U(t). \end{cases}$$

In order to analyze the a posteriori error estimator in Section 5 for time-independent space meshes, we require estimates of the form

$$(4.2)\qquad \left\{ \begin{array}{l} \|\theta(t)\|_E = O(h^{\sigma_1}) \\ \|e_t(t)\|_0 = O(h^{\sigma_2}) \end{array} \right\} \text{ as } h \to 0;\quad \sigma_1 > 1,\ \sigma_2 > 0,$$

with similar estimates needed with $\sigma_2 > 1$ for time-dependent meshes. Estimates of this type for Galerkin approximations of time-dependent equations have been obtained by many authors (cf. [9], [10], [17], [18]), usually as a means of bounding some norm of $\frac{\partial^j}{\partial t^j}e$ for $j \geq 0$, and often where the regularity of the exact solution $u$ is given implicitly by norms of $u$ appearing in the estimates. To clarify the connection with section 5, throughout this section we shall assume for simplicity that

$$(4.3)\qquad P \text{ is a } \kappa(\lambda) \text{ - regular mesh family,}$$

and that our choice of discrete initial data is fixed as

$$(4.4) \qquad U_0^\Delta = \begin{cases} P_1^\Delta u_0; \ \Delta \in \mathcal{P}, & \text{if } u_0 \in \mathcal{D}(L), \\ P_0^\Delta u_0; \ \Delta \in \mathcal{P}, & \text{if } u_0 \notin \mathcal{D}(L). \end{cases}$$

In a manner similar to that used in [9], other choices for $U_0^\Delta$ can be handled, but we require (4.2) to hold uniformly as $t \to 0$ for our smooth data estimates valid on $[0,T)$. Note that for $u_0 \in \mathcal{D}(L)$, $U_t^\Delta(0) = P_0^\Delta u_t(0)$ by (2.9).

The main result of this section is

Theorem 4.1. Let $\alpha,\beta$ satisfy (3.1) and $\alpha \le 4$. Assume (4.3) and (4.4), let $0 \le T_* < T$, and $u(;u_0,f) \in H^{\alpha,\beta}(T_*,T)$. Then $\exists$ a positive constant $C$ (depending on the functions $a$ and $b$, constants $\lambda$ and $T$, but not on $T_*$, $u_0$, $f$, or $u$) such that for each $\Delta \in \mathcal{P}$

$$\left.\begin{aligned} (4.5) \quad \|\theta(t)\|_E &\le C\, h^{\sigma_1}(\Delta) C_\alpha(\beta,T_*,T,u) \\ (4.6) \quad \|e_t(t)\|_0 &\le C\, h^{\sigma_2}(\Delta) C_\alpha(\beta,T_*,T,u) \end{aligned}\right\} \ \forall\ t \in [T_*,T),$$

where

$$(4.7) \qquad \sigma_1 = \min(\alpha,2),$$

$$(4.8) \qquad \sigma_2 = \begin{cases} \alpha-\kappa & , \text{ if } \alpha \le 2, \\ \max(2-\kappa,\alpha-2), & \text{if } 2 < \alpha \le 4. \end{cases}$$

Proof. For each $\alpha$ and $\beta$, let $u_0 = u_0^1 + u_0^2$ be the $\alpha,\beta$ -dependent decomposition of $u_0$ given in (3.18). Then $u = u^1 + u^2 + u^3$ solves eqs. (2.14), where

$$(4.9) \qquad \begin{cases} u^1 = u(;u_0^1,f), \\ u^2 = u(;u_0^2,0), \\ u^3 = 0. \end{cases}$$

Similarly, we decompose the solution $U$ of eqs. (2.17) as $U = U^1 + U^2 + U^3$, where

$$(4.10) \qquad \begin{cases} U^1 = U(;U_0^1,P_0f); & U_0^1 = P_1 u_0^1, \\ U^2 = U(;U_0^2,0); & U_0^2 = P_0 u_0^2, \\ U^3 = U(;U_0^3,0); & U_0^3 = U_0 - U_0^1 - U_0^2. \end{cases}$$

We then have

$$e(t) = \sum_{m=1}^{3} e^m(t); \quad e^m(t) = u^m(t) - U^m(t), \qquad m = 1,2,3,$$

$$\rho(t) = \sum_{m=1}^{3} \rho^m(t); \quad \rho^m(t) = P_1 u^m(t) - u^m(t),$$

$$\theta(t) = \sum_{m=1}^{3} \theta^m(t); \quad \theta^m(t) = P_1 u^m(t) - U^m(t).$$

Using the t-independence of the operators $L$, $L_\Delta$ and subtracting eqs. (2.16) from eqs. (2.15) we get

$$(4.11)\quad \begin{cases} <\theta_t^3(t),\varphi> + B(\theta^3(t),\varphi) \;=\; 0; \quad \forall\ t,\varphi \in (0,T)\times S, \\ \theta^3(0) \;=\; -U_0^3, \end{cases}$$

$$(4.12)\quad B(e^2(t),\varphi) = B(\theta^2(t),\varphi) = -<e_t^2(t),\varphi>; \ \forall\ t,\varphi \in (0,T)\times S,$$

$$(4.13)\quad \begin{cases} < \dfrac{\partial^{j+1}}{\partial t^{j+1}}\theta^1(t),\varphi> + B(\dfrac{\partial^j}{\partial t^j}\theta^1(t),\varphi) \;=\; < \dfrac{\partial^{j+1}}{\partial t^{j+1}}\rho^1(t),\varphi>; \\ \qquad\qquad\qquad\qquad\qquad\qquad \forall\ t,\varphi \in (0,T)\times S, \\ \theta^1(0) \;=\; 0, \qquad\qquad\qquad\qquad j = 0 \ \text{or}\ 1, \\ \theta_t^1(0) \;=\; (P_1-P_0)u_t^1(0). \end{cases}$$

The applications of discrete smoothing Lemma 2.2, spectral representations, and the technique of energy estimates in eqs. (4.11), (4.12), and (4.13), respectively, enable us to demonstrate (4.5) and (4.6) for each of the three components. The triangle inequality then completes the proof of the theorem.

We first note that the choice of $U_0$ in (4.4) and the decompositions (3.18), (4.9), and (4.10) show that for $\beta = \alpha$, $u^2 = u^3 = U^2 = U^3 = 0$. It therefore suffices to show (4.5) and (4.6) for the $m = 2,3$ components only for $\beta < \alpha$.

Specifically, we shall show that for $t \geq T_* > 0$ and $\beta < \alpha$

$$\left.\begin{array}{ll}(4.14) & \|\theta^3(t)\|_E \\ (4.15) & \|e_t^3(t)\|_0\end{array}\right\} \le C(1+T_*^{-(\alpha+1)/2})\|U_0^3\|_0,$$

$$(4.16) \qquad \|\theta^2(t)\|_E \le C\, h^{\sigma_1}(1+T_*^{-(\alpha+1)/2})\|u_0^2\|_0,$$

$$(4.17) \qquad \|e_t^2(t)\|_0 \le C\, h^{\sigma_2}(1+T_*^{-(\alpha+1)/2})\|u_0^2\|_0,$$

and for $0 \le t < T$ and $\beta \le \alpha$

$$(4.18) \qquad \|\theta^1(t)\|_E \le C\, h^{\sigma_1}\|u^1\|^{(0,T)}_{\alpha+2,(\alpha+2)/2},$$

$$(4.19) \qquad \|e_t^1(t)\|_0 \le C\, h^{\sigma_2}\|u^1\|^{(0,T)}_{\alpha+2,(\alpha+2)/2}.$$

It follows from (3.18), (3.20), (4.4), (4.7), (4.8), (4.10), Lemma 2.1, and the triangle inequality that for $\beta < 1 < \alpha$ (i.e. $u_0 \notin \mathcal{D}(L)$)

$$(4.20) \qquad \|U_0^3\|_0 = \|(P_0-P_1)\psi_\alpha\|_0 \le C\, h^2\|\psi_\alpha\|_{\alpha+1} \le C \cdot \min(h^{\sigma_1},h^{\sigma_2}) \cdot M_\alpha^{CR}(0,f),$$

and if $1 \le \beta < \alpha$ (i.e. $u_0 \in \mathcal{D}(L)$)

$$(4.21) \qquad \|U_0^3\|_0 = \|(P_1-P_0)u_0 + (P_0-P_1)\psi_\alpha\|_0 \le C\, h^2\{\|u_0\|_2 + \|\psi_\alpha\|_2\}$$
$$\le C \cdot \min(h^{\sigma_1},h^{\sigma_2}) \cdot \{\|u_0\|_{\beta+1} + M_\alpha^{CR}(0,f)\}.$$

Also, since $u^1 \in H^{\alpha+2,(\alpha+2)/2}(0,T)$ for $\beta \le \alpha$, we have by Thm. 3.2 that

$$(4.22) \qquad \|u^1\|^{(0,T)}_{\alpha+2,(\alpha+2)/2} \le C\{\|u^1_0\|_{\alpha+1} + \|f\|^{(0,T)}_{\alpha,\alpha/2}\}.$$

Using (4.20)-(4.22), the estimates (3.14) for $u^1_0$ and $u^2_0$, and the definition (3.9) of $C_I$ in (4.14)-(4.19), we see that the demonstration of (4.14)-(4.19) will complete the proof of Thm. 4.1.

We shall make use of the following simple lemma.

<u>Lemma 4.1</u>. Let $0 \le \gamma_1 \le \gamma_2$ and $0 < T_* \le t$. Then $t^{-\gamma_1} \le 1 + T_*^{-\gamma_2}$.

We first demonstrate (4.14) and (4.15). By (4.9), (4.10), eqs. (4.11), and Lemmas 2.2 and 4.1 we have that for $\alpha > 1$ and $t \ge T_* > 0$

$$(4.23) \qquad \|\theta^3(t)\|_E = \|U^3(t)\|_E \le C\, t^{-1/2}\|U^3_0\|_0 \le C(1+T_*^{-(\alpha+1)/2})\|U^3_0\|_0,$$

$$(4.24) \qquad \|e^3_t(t)\|_0 = \|\frac{\partial}{\partial t} U^3(t)\|_0 \le C\, t^{-1}\|U^3_0\|_0 \le C(1+T_*^{-(\alpha+1)/2})\|U^3_0\|_0,$$

which are (4.14) and (4.15).

Because $e^2$ satisfies eqs. (4.12) and $U^2_0 = P_0u^2_0$, spectral representations may be applied to obtain (4.16) and (4.17). A simple interpolation argument leads to the following slight variation of [9, Thm. 3.1].

Theorem 4.2. There exists a positive constant $C$ such that for $j \geq 0$, $0 \leq s \leq 2$, and $t \geq T_* > 0$

$$(4.25) \qquad \left\|\frac{\partial^j}{\partial t^j} e^2(t)\right\|_0 \leq C\, h^s\, T_*^{-(s+2j)/2} \|u_0^2\|_0 .$$

It is easily verified in (4.8) that

$$\sigma_2 \leq \begin{cases} \alpha - 1 \in (0,2), & \text{if } \alpha \in (1,3), \\ 2, & \text{if } \alpha \in (3,4]. \end{cases}$$

The application of (4.25) with $j = 1$ and $s = \alpha - 1$ or $2$ yields

$$(4.26) \qquad \|e_t^2(t)\|_0 \leq C\, h^{\alpha-1}\, T_*^{-(\alpha+1)/2} \|u_0^2\|_0; \quad \text{if } 1 < \alpha < 3, \text{ and}$$

$$(4.27) \qquad \|e_t^2(t)\|_0 \leq C\, h^2\, T_*^{-2} \|u_0^2\|_0; \qquad \text{if } 3 < \alpha \leq 4.$$

The estimate (4.17) then follows by applying Lemma 4.1.

To show (4.16) we set $\varphi = \theta^2 = \rho^2 + e^2$ in eq. (4.12) and get for $t \geq T_* > 0$ and any $s \in [0,2]$, $s \neq 1/2$

$$\|\theta^2(t)\|_E^2 = |< e_t^2(t), \rho^2(t) + e^2(t)>|$$

by the triangle inequality, (4.1), and Lemma 2.1

$$\leq \|e_t^2(t)\|_0 \{C\, h^s \|u^2(t)\|_s + \|e^2(t)\|_0\}$$

by Lemma 3.1 and Thm. 4.2

$$\leq C\, h^{2s}\, t^{-(s+1)} \|u_0^2\|_0^2, \quad \text{and hence}$$

$$(4.28) \qquad \|\theta^2(t)\|_E \le C\, h^s\, T_*^{-(s+1)/2} \|u_0^2\|_0 .$$

Setting $s = \sigma_1 = \min(\alpha,2)$ for $1 < \alpha \le 4$ and using Lemma 4.1 we get (4.16).

We proceed to demonstrate (4.18) and (4.19) directly via energy estimates. These results can also be obtained for some, but not all $\alpha \in (1,4]$ satisfying (3.1) by applying Thms. 2.1, 2.2, and results given in [17] in arguments similar to (4.25)-(4.28).

Setting $\varphi = 2\theta_s^1(s)$ in eqs. (4.13) for $j = 0$ and $0 < s \le t < T$, we get

$$(4.29) \quad 2\|\theta_s^1(s)\|_0^2 + 2B(\theta^1(s),\ \theta_s^1(s)) = 2\langle\rho_s^1(s),\theta_s^1(s)\rangle .$$

Using the $t$-independence of the symmetric form $B$ and the fact that $\theta^1(0) = 0$, integration of eq. (4.29) on $(0,t)$ followed by the application of the Cauchy-Schwarz inequality yields

$$(4.30) \qquad \|\theta^1(t)\|_E^2 \le \frac{1}{2}\int_0^t \|(I-P_1)u_s^1(s)\|_0^2\, ds,$$

from which we obtain

$$(4.31) \qquad \|\theta^1(t)\|_E \le C\|(I-P_1)u_t^1\|_{0,0}^{(0,T)} .$$

For $1 < \alpha \le 2$, $u^1 \in H^{\alpha+2,(\alpha+2)/2}(0,T) \cap H^0((0,T);H_0^1)$ and so by Lemma 2.3 and (4.7)

$$(4.32) \qquad \|\theta^1(t)\|_E \le C\, h^{\sigma_1} \|u^1\|^{(0,T)}_{\alpha+2,(\alpha+2)/2},$$

which is (4.18) for $1 < \alpha \le 2$. The result (4.18) for $2 < \alpha \le 4$ follows trivially.

It remains to demonstrate (4.19). For $1 < \alpha \le 4$, by (2.9) and (4.10) we have that for $0 \le t < T$

$$(4.33) \qquad e^1_t(t) = e^{11}_t(t) + e^{12}_t(t), \text{ where}$$

$$(4.34) \qquad e^{11}_t(t) = (I-P_0)u^1_t(t), \text{ and}$$

$$(4.35) \qquad e^{12}_t(t) = L_\Delta \theta^1(t).$$

By Lemma 2.1, Thm. 2.2, and $\kappa \ge 1$, we have that for $1 < \alpha \le 2$, $\alpha \ne 3/2$

$$(4.36) \quad \|e^{11}_t(t)\|_0 \le C\, h^{\alpha-\kappa} \sup_{s\in[0,T]} \|u^1_t(s)\|_{\alpha-1} \le C\, h^{\alpha-\kappa} \|u^1\|^{(0,T)}_{\alpha+2,(\alpha+2)/2},$$

and similarly for $2 < \alpha \le 4$ with $h^{\alpha-\kappa}$ replaced by $h^{2-\kappa}$. By (2.11) and (4.18) we also have that for $1 < \alpha \le 4$

$$(4.37) \quad \|e^{12}_t(t)\|_0 = \|L_\Delta \theta^1(t)\|_0 \le C\, h^{-\kappa} \|\theta^1(t)\|_E \le C\, h^{\sigma_1-\kappa} \|u^1\|^{(0,T)}_{\alpha+2,(\alpha+2)/2}.$$

By the triangle inequality and the definition (4.8) of $\sigma_2$, we see that to complete the proof of (4.19), it suffices to show

$$(4.38) \quad \|e_t^1(t)\|_0 \le C\, h^{\alpha-2} \|u^1\|_{\alpha+2,(\alpha+2)/2}^{(0,T)}, \quad \text{for} \quad 2 < \alpha \le 4.$$

We have that $e_t^1(t) = (I-P_1)u_t^1(t) + \theta_t^1(t)$.

It follows as in (4.36) that

$$(4.39) \qquad \|(I-P_1)u_t^1(t)\|_0 \le C\, h^{\alpha-2} \|u^1\|_{\alpha+2,(\alpha+2)/2}^{(0,T)}.$$

Setting $\varphi = 2\theta_s^1(s)$ in eqs. (4.13) for $j = 1$ and $0 < s \le t < T$, we get

$$(4.40) \quad \frac{d}{ds}\|\theta_s^1(s)\|_0^2 + 2\|\theta_s^1(s)\|_E^2 = 2<\rho_{ss}^1(s),\ \theta_s^1(s)>.$$

Integration of (4.40) on $(0,t)$ followed by the use of the *Cauchy-Schwarz inequality* yields

$$(4.41) \quad \|\theta_t^1(t)\|_0^2 \le \|\theta_t^1(0)\|_0^2 + C\int_0^t \|(I-P_1)u_{ss}^1(s)\|_0^2\, ds.$$

Using $\theta_t^1(0) = (P_1-P_0)u_t^1(0)$, we get

$$(4.42) \quad \|\theta_t^1(t)\|_0 \le C\{\|(P_1-I)u_t^1(0)\|_0 + \|(I-P_0)u_t^1(0)\|_0 + \|(I-P_1)u_{tt}^1\|_{0,0}^{(0,T)}\}.$$

Using Lemmas 2.1, 2.3, and Thm. 2.2 we get from (4.42) that

$$(4.43) \quad \|\theta_t^1(t)\|_0 \le C\, h^{\alpha-2}\|u^1\|_{\alpha+2,(\alpha+2)/2}^{(0,T)} \quad \text{for} \quad 2 < \alpha \le 4.$$

The proof of (4.38) and Thm. 4.1 is then completed by using (4.39), (4.43), and the triangle inequality.

<u>Remarks</u>. The following consequences Thm. 4.1 are relevant to the estimates in section 5. Define the function $\bar{\kappa}$ by

$$(4.44)\qquad \bar{\kappa}(\alpha) = \begin{cases} (\alpha+1)/2; & 1 < \alpha \le 2, \\ 3/2 ; & 2 \le \alpha \le 3, \\ 2 ; & 3 < \alpha \le 4. \end{cases}$$

For fixed $\alpha \in (1,4]$ and $\kappa \in [1,\bar{\kappa}(\alpha))$ set

$$(4.45)\qquad \mu(\kappa,\alpha) = \frac{\bar{\kappa}(\alpha)-\kappa}{2}.$$

It can be verified from the result of Thm. 4.1 that $\exists$ a positive constant $C$ such that for each $\Delta \in P$ and all $t \in [T_*,T)$

$$\left.\begin{array}{ll} (4.46) & h^{1-\kappa}(\Delta)\{\|\theta(t)\|_E + \|e_t(t)\|_0\} \\ (4.47) & h^{-\kappa}(\Delta)\|\theta(t)\|_E \end{array}\right\} \le C\, h^{2\mu}(\Delta) C_\alpha(\beta,T_*,T,u).$$

## 5. A Posteriori Error Estimates

All notation used in Section 4 is adopted throughout this section. The solution classes $\mathcal{H}^{\alpha,\beta}_{\delta_1}(T_*,T)$ and $\mathcal{H}^{\alpha,\beta}_{\delta_1,\delta_2}(T_*,T)$ are as defined in (3.21) - (3.27).

We begin by listing for easy reference the assumptions concerning solution regularity, mesh family type, and discrete initial data choice which will lead to our results.

$$(5.1)\quad \begin{cases} u \in \mathcal{H}^{\alpha,\beta}(T_*,T); \quad 0 \le T_* < T, \quad \alpha \text{ and } \beta \text{ satisfy} \\ (3.1) \text{ and } \alpha \le 4. \text{ For } \Delta \in \mathcal{P}, \ U_0^\Delta \text{ is chosen according} \\ \text{to } (4.4). \end{cases}$$

$$(5.2)\quad \begin{cases} u \in \mathcal{H}^{\alpha,\beta}_{\delta_1}(T_*,T); \quad \delta_1 > 0. \ \mathcal{P} \text{ is a } \kappa(\lambda)\text{-regular} \\ \text{mesh family.} \end{cases}$$

$$(5.3)\quad u \in \mathcal{H}^{\alpha,\beta}_{\delta_1,\delta_2}(T_*,T); \quad \alpha > 3, \quad \delta_1 > 0, \quad \delta_2 > 0.$$

$$(5.4)\quad \kappa \in [1,\bar{\kappa}(\alpha)); \quad \bar{\kappa}(\cdot) \text{ defined as in } (4.44).$$

Whenever they appear, unless otherwise specified

(5.5) $C, C_1, C_2$, etc. denote positive generic constants which depend on the functions $a$ and $b$, constants $\lambda$ and $T$, but not on $T_*, \alpha, \beta, \delta_1, \delta_2, u, U$, or the data $u_0$ and $f$.

We now define certain concepts needed in our a posteriori analysis which are related to those developed for elliptic boundary value problems in [2], [3], [4], and [5].

For $0 \le T_* < T$ let

$$\left.\begin{aligned} &(5.6) \qquad g:\ \Delta \to g^{\Delta} \in C^0([T_*,T);\ H^0) \\ &(5.7) \qquad \phi:\ \Delta \to \phi^{\Delta} \in C^0([T_*,T);\ S(\Delta)) \end{aligned}\right\}\ \Delta \in P.$$

We define the local <u>indicators</u>

$$(5.8) \qquad \eta_j(t,g^{\Delta},\phi^{\Delta}) = \frac{h_j}{(12a_j)^{1/2}} \|L\phi^{\Delta}(t) - g^{\Delta}(t)\|_{0,I_j};$$

$$t \in [T_*,T), \quad 1 \le j \le N(\Delta),$$

and the associated <u>estimator</u>

$$(5.9) \qquad E(t,g^{\Delta},\phi^{\Delta}) = \left\{\sum_{j=1}^{N(\Delta)} \eta_j^2(t,g^{\Delta},\phi^{\Delta})\right\}^{1/2};$$

$$t \in [T_*,T).$$

By (5.7),

$$(5.10) \quad L\phi^{\Delta}(t,x) = -a_x(x)\phi_x^{\Delta}(t,x) + b(x)\phi^{\Delta}(t,x);$$

$$t,x \in [T_*,T) \times I_j, \quad 1 \le j \le N(\Delta),$$

and therefore (5.8) and (5.9) are well-defined.

Let $\Phi \in C^0([T_*,T);\ H_0^1)$ and $g,\phi$ be as in (5.6) and (5.7). We say that a quantity $E(\cdot,g,\phi)$ is an <u>upper estimator</u> for $\|\Phi - \phi\|_E$ on $[T_*,T)$ if $\exists$ positive constants

$C_1, C_2, h^*$ such that for $\Delta \in P$ with $h(\Delta) \le C_1 h^*$ and all $t \in [T_*, T)$

$$(5.11) \qquad \|\Phi(t) - \phi^\Delta(t)\|_E \le C_2\, E(t, g^\Delta, \phi^\Delta),$$

where $C_1$ and $C_2$ are as in (5.5) and $h^*$ depends in general on the mesh family $P$ {i.e.g. on $\kappa$} and on the class in which $\Phi$ resides {i.e.g. on $\alpha, \beta, \delta_1, \delta_2$}.

If under the same hypotheses $\exists$ a constant $\sigma > 0$ such that

$$(5.12) \quad \|\Phi(t) - \phi^\Delta(t)\|_E = E(t, g^\Delta, \phi^\Delta)(1 + O(h^\sigma)) \quad \text{as} \quad h \to 0,$$

where the constant in the $O$-term is as in (5.5), then we say that $E(\cdot, g, \phi)$ is an <u>asymptotically exact estimator</u> for $\|\Phi - \phi\|_E$ on $[T_*, T)$.

Let $u$ and $U$ be the solutions of eqs.(2.14) and (2.17). We take as our primary estimator for the error $\|u - U\|_E$ the quantity $E(\cdot, f - U_t, U)$, defined according to (5.8) and (5.9). The principal results in this section are the following two theorems.

<u>Theorem 5.1</u>. Assume (5.1), (5.2), (5.4), and let $\mu = \mu(\kappa, \alpha)$ be as defined in (4.45). Then $E(\cdot, f\text{-}U_t, U)$ is an upper estimator for $\|u\text{-}U\|_E$ on $[T_*, T)$, with $h^* = \delta_1^{-1/\mu}$.

Theorem 5.2. Assume (5.1) - (5.4) and let $\mu$ be as defined in (4.45). Then $\mathcal{E}(\cdot, f-U_t, U)$ is an asymptotically exact estimator for $\|u-U\|_E$ on $[T_*, T)$, with $h^* = \min(\delta_1^{-1/\mu}, \delta_2^{-1/\mu})$ and $\sigma = \mu$.

$\mathcal{E}(\cdot, f - U_t, U)$ is based upon local, computable residuals and can therefore be used to monitor and control the space discretization error in an adaptive FEMOL procedure. While $h^*$ may be quite small, its existence indicates that $\mathcal{E}(\cdot, f - U_t, U)$ will perform well as an estimator for a given class of data, and not be overly sensitive to changes in solution behavior during the time evolution of a problem.

The proof of Thms. 5.1, 5.2 is based upon interpreting $u_t(t)$ as given data, eqs. (2.14) as a continuous, one-parameter (t) family of elliptic boundary value problems:

For each $t \in [T_*, T)$, find $u(t) \in \mathcal{D}(L)$ satisfying

$$Lu(t) = f(t) - u_t(t), \tag{5.13}$$

and employing the error $\|\rho(t)\|_E$ and estimator $\mathcal{E}(t, f - u_t, P_1 u)$ associated with (5.13). $\mathcal{E}(\cdot, f - u_t, P_1 u)$ is defined according to (5.8) and (5.9) and is not computable, and $\|\rho\|_E$ is not of primary interest, but up to higher order terms in $h$ we have $\mathcal{E}(\cdot, f - u_t, P_1 u) \sim \mathcal{E}(\cdot, f - U_t, U)$ and $\|\rho\|_E \sim \|e\|_E$.

Specifically, the proofs of Thms. 5.1 and 5.2 follow from the application of Thm. 4.1 and the demonstration of the following sequence of lemmas.

<u>Lemma 5.1</u>. There exists a positive constant $C$ such that for each $\Delta \in P$ and all $t \in [T_*,T)$

$$|E(t,f-U_t^\Delta,U^\Delta) - E(t,f-u_t,P_1^\Delta u)| \le Ch(\Delta)\{\|\theta(t)\|_E + \|e_t(t)\|_0\}.$$

<u>Lemma 5.2</u>. Assume (5.1). Then $E(\cdot,f-u_t,P_1 u)$ is an upper estimator for $\|\rho\|_E$ on $[T_*,T)$, with $h^* = 1$.

<u>Lemma 5.3</u>. Assume (5.1), (5.2), and the result of Lemma 5.2. Then $\exists$ a positive constant $C$ such that for each $\Delta \in P$ and all $t \in [T_*,T)$

$$E(t,f-u_t,P_1^\Delta u) \ge \frac{Ch^\kappa(\Delta)}{\delta_1} C_\alpha(\beta,T_*,T,u).$$

<u>Lemma 5.4</u>. Assume (5.1) - (5.4). Then $E(\cdot,f-u_t,P_1 u)$ is an asymptotically exact estimator for $\|\rho\|_E$ on $[T_*,T)$, with $h^* = \delta_2^{-1/\mu}$, $\mu$ defined as in (4.45), and $\sigma = \mu$.

Before proving Lemmas 5.1 - 5.4, we show that the result of Thm. 5.1 follows from the validity of hypotheses (5.1), (5.2), (5.4), and the results of Lemmas 5.1, 5.2, 5.3, and, if the result of Lemma 5.4 also holds, then we have the conclusion of Thm. 5.2.

To simplify notation we suppress $t$, and also write

$$
(5.14)\qquad \begin{cases} \mathcal{E}_1 = \mathcal{E}(t, f - U_t^{\Delta}, U^{\Delta}), \\ \mathcal{E}_2 = \mathcal{E}(t, f - u_t, P_1^{\Delta} u), \\ C_\alpha = C_\alpha(\beta, T_*, T, u). \end{cases}
$$

By Lemma 5.3 we may write

$$
(5.15)\qquad 1 - \frac{|\mathcal{E}_1 - \mathcal{E}_2|}{\mathcal{E}_2} \le \frac{\mathcal{E}_1}{\mathcal{E}_2} \le 1 + \frac{|\mathcal{E}_1 - \mathcal{E}_2|}{\mathcal{E}_2},
$$

which after application of Lemma 5.1 and, again, Lemma 5.3 becomes

$$
(5.16)\qquad 1 - \frac{C\delta_1 h^{1-\kappa}}{C_\alpha} \{\|\theta\|_E + \|e_t\|_0\} \le \frac{\mathcal{E}_1}{\mathcal{E}_2} \le 1 + \frac{C\delta_1 h^{1-\kappa}}{C_\alpha} \{\|\theta\|_E + \|e_t\|_0\}.
$$

Under assumptions (5.1), (5.2), and (5.4), we may apply the consequences (4.44) - (4.47) of Thm. 4.1 to (5.16) and get

$$
(5.17)\qquad 1 - C\delta_1 h^{2\mu} \le \frac{\mathcal{E}_1}{\mathcal{E}_2} \le 1 + C\delta_1 h^{2\mu}.
$$

Now, by the definition of $P_1 : H_0^1 \to S$

$$
(5.18)\qquad \|\rho\|_E = \inf_{\psi \in S} \|u - \psi\|_E \le \|e\|_E,
$$

and so by applying the triangle inequality to $e = -\rho + \theta$ we have

$$(5.19) \qquad \|\rho\|_E \le \|e\|_E \le \|\rho\|_E + \|\theta\|_E .$$

By Lemma 5.3, (5.19) becomes

$$(5.20) \qquad \mathcal{E}_2 \frac{\|\rho\|_E}{\mathcal{E}_2} \le \|e\|_E \le \mathcal{E}_2 \left\{ \frac{\|\rho\|_E}{\mathcal{E}_2} + \frac{C\delta_1 h^{-\kappa}}{C_\alpha} \|\theta\|_E \right\},$$

and using (4.44) - (4.47) as in (5.17)

$$(5.21) \qquad \mathcal{E}_2 \frac{\|\rho\|_E}{\mathcal{E}_2} \le \|e\|_E \le \mathcal{E}_2 \left\{ \frac{\|\rho\|_E}{\mathcal{E}_2} + C\delta_1 h^{2\mu} \right\}.$$

Setting $h^* = \delta_1^{-1/\mu}$, we see in (5.17) that $\exists$ a positive constant $C_1$ such that for $h \le C_1 h^*$,

$$(5.22) \qquad \mathcal{E}_2 = \mathcal{E}_1(1 + O(h^\mu)) \quad \text{as} \quad h \to 0,$$

with constants in $O(h^\mu)$ as in (5.5), and therefore from (5.21) we get

$$(5.23) \quad \|e\|_E = \mathcal{E}_1 \frac{\|\rho\|_E}{\mathcal{E}_2} (1 + O(h^\mu)) \quad \text{as} \quad h \to 0.$$

The demonstration of Thms. 5.1, 5.2 now follows easily from the results of Lemmas 5.2, 5.4, respectively.

We proceed to prove Lemmas 5.1 - 5.4.

<u>Proof of Lemma 5.1</u>. To further simplify notation we write $u^\Delta = P_1^\Delta u$ and for $1 \le j \le N(\Delta)$, $\eta_j(U) = \eta_j(t,f-U_t^\Delta,U^\Delta)$ and $\eta_j(u^\Delta) = \eta_j(t,f-u_t,u^\Delta)$. By definitions (5.8), (5.9), and the triangle inequality we have

$$(5.24) \qquad |\varepsilon_1^2-\varepsilon_2^2| \le \sum_{j=1}^{N} |\eta_j^2(U)-\eta_j^2(u^\Delta)|$$

$$\le \sum_{j=1}^{N} \frac{h_j^2}{12a_j} |Q_{1j}-Q_{2j}| \cdot |Q_{1j}+Q_{2j}|,$$

where

$$\left.\begin{aligned} Q_{1j} &= \|LU + U_t - f\|_{0,I_j} \\ Q_{2j} &= \|Lu^\Delta + u_t - f\|_{0,I_j} \end{aligned}\right\} \quad 1 \le j \le N.$$

Using again the triangle inequality, the $N$-dimensional Cauchy-Schwarz inequality, and (5.8) we get from (5.24) that

$$(5.25) \qquad |\varepsilon_1^2-\varepsilon_2^2| \le 2\left\{\sum_{j=1}^{N} \frac{h_j^2}{12a_j} (\|L\theta\|_{0,I_j}^2+\|e_t\|_{0,I_j}^2)\right\}^{1/2}$$

$$\cdot \left\{\sum_{j=1}^{N}(\eta_j^2(U)+\eta_j^2(u^\Delta))\right\}^{1/2}.$$

It then follows from (2.1), (2.2), (5.9), and (5.10) that $\exists$ a positive constant $C$ such that

$$(5.26)\quad |E_1^2 - E_2^2| \le Ch(\Delta)\{\|\theta\|_E + \|e_t\|_0\}\{E_1 + E_2\},$$

from which the desired result follows.

Proof of Lemma 5.2. By Thms. 2.2, 2.3, and 3.1, the assumption (5.1) guarantees that $u \in C^0([T_*,T);\ \mathcal{D}(L))$. Interpreting eqs. (2.14) as in (5.13) and applying [2, Thm. 7.1] (stated for $u_{xx} \in C^0(I)$, but valid in fact for $u_{xx} \in H^0$; cf. also [4]), we have the desired result.

*Proof of Lemma 5.3*. Letting again $u^\Delta = P_1^\Delta u$, we note that for $t,x \in [T_*,T) \times I_j$; $1 \le j \le N$

$$(5.27)\quad Lu^\Delta(t,x) - (f(t,x) - u_t(t,x))$$

$$= (u_t - f - a_x u_x^\Delta + bu^\Delta)(t,x) = (au_{xx} - a_x\rho_x + b\rho)(t,x).$$

It then follows from (5.8) that $\exists$ a positive constant $C$ such that

$$(5.28)\quad \eta_j^2(t, f - u_t, u^\Delta) \ge Ch_j^2\{\|u_{xx}(t,\cdot)\|_{0,I_j}^2 - \|\rho(t)\|_{E,I_j}^2\};$$

$$1 \le j \le N.$$

Summing for $j = 1,N$ yields

$$(5.29)\quad E^2(t, f - u_t, u^\Delta) + Ch^2(\Delta)\|\rho(t)\|_E^2 \ge C\underline{h}^2(\Delta)\|u_{xx}(t,\cdot)\|_0^2,$$

which by the $\kappa$-regularity of assumption (5.2) and the result of Lemma 5.2 becomes

$$(5.30) \quad \mathcal{E}^2(t, f - u_t, u^\Delta) \geq \left(\frac{C}{1+Ch^2(\Delta)}\right) h^{2\kappa}(\Delta) \|u_{xx}(t,\cdot)\|_0^2 .$$

Since by (5.2), $\|u_{xx}(t,\cdot)\|_0 \geq \frac{\mathcal{C}_\alpha}{\delta_1}$; $\forall\, t \in [T_*, T)$, we have from (5.30) the result of Lemma 5.3.

We omit the proof of Lemma 5.4, as it is completely analogous to that in [2, Thm. 7.3], where a sequence of six asymptotic equalities related to a decomposition of the indicators led to the desired result (cf. [8] for further details). Roughly speaking, the properties (3.24) - (3.27) assumed for asymptotic exactness require that the solution is smooth and not relatively flat at any time $t \in [T_*, T)$.

## 6. A Posteriori Error Estimates Revisited

The purpose in using a posteriori, rather than a priori estimates to bound the error due to the numerical solution of differential equations is to obtain more realistic bounds. We expect that restrictive assumptions on admissible solution classes and mesh partitions need to be imposed in order to guarantee that an estimator is asymptotically exact. However, it is desirable that an estimator is reliable, i.e. an upper estimator, under far less restrictive and more easily verifiable conditions.

We shall show that if the estimator $\mathcal{E}(\cdot, f - U_t, U)$ introduced in Section 5 is modified through the addition of a term, which takes more fully into account the t-dependence of the problem and is of higher order in $h$, then we obtain a reliable estimator for a class of problems occurring often in practice, with no restrictions on the mesh family or mesh size.

We assume that the solutions $u(\ ;u_0,f)$ and $U(\ ;U_0,P_0f)$ of eqs. (2.15), (2.16) are such that

$$f, f_t \in C^0([0,T);H^0), \tag{6.1}$$

$$u_0 \in \mathcal{D}(L), \quad \text{and} \tag{6.2}$$

$$U_0^\Delta = P_1^\Delta u_0, \quad \text{for} \quad \Delta \in \mathcal{P}. \tag{6.3}$$

Note that from (6.1) and (6.2) we have that

$$u \in C^0([0,T); \mathcal{D}(L)) \quad \text{and} \quad u_t \in C^0([0,T); H^0).$$

Define $\hat{C} = \underline{a}\pi^2$ ($\underline{a}$ as in (2.1)) $\leq \lambda_1 \equiv \inf\limits_{\substack{\phi \in H_0^1 \\ \phi \neq 0}} \dfrac{\|\phi\|_E^2}{\|\phi\|_0^2}$ .

For $t \in [0,T)$ and $\Delta \in P$ we define the estimator

$$\text{(6.4)} \qquad E(t, f - U_t^\Delta, U^\Delta) = \sum_{\ell=1}^{3} E_\ell(t,\Delta), \quad \text{where}$$

$$\text{(6.5)} \qquad E_1(t,\Delta) = \mathcal{E}(t, f - U_t^\Delta, U^\Delta)$$

$$\text{(6.6)} \qquad E_2(t,\Delta) = h(\Delta)\left\{\int_0^t e^{-\hat{C}(t-s)}\mathcal{E}^2(s, f_s - U_{ss}^\Delta, U_s^\Delta)ds\right\}^{1/2},$$

(6.7)

$$E_3(t,\Delta) = \begin{cases} e^{-\hat{C}\,t/2} \cdot \min(1, h(\Delta)/t^{1/2}) \cdot \mathcal{E}(0, U_t^\Delta(0) + Lu_0 - f(0), 0) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad , \text{ if } t > 0, \\ \mathcal{E}(0, U_t^\Delta(0) + Lu_0 - f(0), 0), \quad \text{if} \quad t = 0. \end{cases}$$

By (6.1), (6.2), and the fact that $U^\Delta \in C^\infty((0,T); S(\Delta))$ we see that $\mathcal{E}(t, f - U_t^\Delta, U^\Delta)$ is well-defined and computable, given $U^\Delta$. Note that

(6.8)

$$E_2(t,\Delta) = h(\Delta)\left\{\sum_{j=1}^{N(\Delta)} \frac{h_j^2}{12a_j} \int_0^t e^{-\hat{C}(t-s)} \left\| \frac{\partial}{\partial s}(U_s^\Delta + LU^\Delta - f)(s,\cdot)\right\|_{0,I_j}^2 ds\right\}^{1/2}$$

$= O(h^2(\Delta))$ as $h(\Delta) \to 0$, provided that $u$ is sufficiently smooth ,

and that by (2.9), (6.1) - (6.3), and Lemma 2.1

(6.9) $$E(0,U_t^\Delta(0)+Lu_0-f(0),0)$$

$$= \{\sum_{j=1}^{N(\Delta)} \frac{h_j^2}{12a_j} \|(I-P_0^\Delta)(Lu_0-f(0))\|_{0,I_j}^2\}^{1/2}$$

$$\le Ch(\Delta)\{\|u_0\|_2 + \|f(0)\|_0\}.$$

Suppose that for $\varepsilon > 0$ arbitrarily small, we either assume

(6.10) in addition to (6.1), (6.2) that $u_0 \in H^{2+\varepsilon}$ and $f(0) \in H^\varepsilon$, or that

(6.11) $$t \ge T_* > 0 \quad \text{and} \quad h(\Delta) \le T_*^{1/2(1-\varepsilon)}.$$

Then from (6.5), (6.7) - (6.9) we see that

(6.12) $E(t,f-U_t^\Delta,U^\Delta) = E(t,f-U_t^\Delta,U^\Delta) + O(h^{1+\varepsilon}(\Delta))$ as $h(\Delta) \to 0$.

Since the exponential decay rates in $E_2(t,\Delta)$ and $E_3(t,\Delta)$ are overly pessimistic for any given problem, and since $E_2(t,\Delta)$ is related to the t-dependence of the local indicators $\{\eta_j\}_{j=1,N(\Delta)}$, a practical strategy for the implementation of $E(t,f-U_t,U)$ would be to use only $E(t,f-U_t,U)$, and to monitor some measure of $\{\dot{\eta}_j\}_{j=1,N(\Delta)}$ on a time interval $[t-C_0h^{\gamma},t]$, for some positive constants $C_0$ and $\gamma$. We give no details here.

We shall show the following result.

<u>Theorem 6.1</u>. Assume (6.1) - (6.3). Then $E(\cdot,f-U_t,U)$ is an upper estimator for $\|u-U\|_E$ on $[0,T)$, with $h^* = 1$.

<u>Proof</u>. The proof is based on a different decomposition of the error $e = u - U$ than that used in Section 5. We write

$$(6.13) \qquad e(t) = e^1(t) + e^2(t), \quad \text{where}$$

$$(6.14) \qquad e^1(t) = v^{\Delta}(t) - U^{\Delta}(t),$$

$$(6.15) \qquad e^2(t) = u(t) = v^{\Delta}(t),$$

and $v^{\Delta} \in C^0([0,T); \mathcal{D}(L))$ is defined by

$$(6.16) \qquad Lv^{\Delta}(t) = (f-U_t^{\Delta})(t); \quad t \in [0,T).$$

By (6.1) and the t-independence of the operator $L$, we have that

$$v_t^{\Delta} \in C^0([0,T); \mathcal{D}(L))$$

and

(6.17) $$Lv_t^\Delta(t) = (f_t - U_{tt}^\Delta)(t); \quad t \in [0,T).$$

From eqs. (2.15), (2.16), and the definition of $P_1^\Delta: H_0^1 \to S(\Delta)$ it follows that

(6.18) $$P_1^\Delta \frac{\partial^j}{\partial t^j} v(t,\cdot) = \frac{\partial^j}{\partial t^j} U^\Delta(t,\cdot); \quad t \in [0,T), \quad j = 0 \text{ or } 1.$$

Subtraction of eqs. (2.16) from eqs. (2.15) shows that

(6.19) $$e^2(t) = e^{21}(t) + e^{22}(t), \quad \text{where}$$

(6.20) $$\begin{cases} \langle e_s^{21}(s),\phi\rangle + B(e^{21}(s),\phi) = -\langle e_s^1(s),\phi\rangle; \forall s,\phi \in (0,T) \times H_0^1, \\ \\ e^{21}(0) = 0, \end{cases}$$

(6.21) $$\begin{cases} \langle e_s^{22}(s),\phi\rangle + B(e^{22}(s),\phi) = 0; \quad \forall\, s,\phi \in (0,T) \times H_0^1, \\ \\ e^{22}(0) = e^2(0). \end{cases}$$

We show that there exists a positive constant $C$ such that for all $t \in [0,T)$ and each $\Delta \in P$

(6.22) $$\|e^1(t)\|_E \le C\, E_1(t,\Delta),$$

(6.23) $$\|e^{21}(t)\|_E \le C\, E_2(t,\Delta), \quad \text{and}$$

(6.24) $$\|e^{22}(t)\|_E \le C\, E_3(t,\Delta),$$

which will complete the proof of Thm. 6.1. By (6.16) and (6.18), the estimate (6.22) follows from the result in

[2, Thm. 7.1], as in the proof of Lemma 5.2.

From the t-independence of the operator $L$ and eqs. (6.20), we get that the function $w(s) = e^{\lambda_1 s/2} e^{21}(s)$ satisfies

(6.25)

$$\begin{cases} \langle w_s(s),\phi\rangle + B(w(s),\phi) - \dfrac{\lambda_1}{2} \langle w(s),\phi\rangle = -e^{\lambda_1 s/2} \langle e^1_s(s),\phi\rangle; \\ \qquad\qquad \forall s,\phi \in (0,T) \times H^1_0, \\ w(0) = 0 . \end{cases}$$

Setting $\phi = 2w_s(s)$ in eq. (6.25) yields

$$(6.26) \quad 2\|w_s(s)\|_0^2 + \frac{d}{ds} \{\|w(s)\|_E^2 - \frac{\lambda_1}{2} \|w(s)\|_0^2\} = -2e^{\lambda_1 s/2} \langle e^1_s(s),w_s(s)\rangle.$$

Integrating (6.26) on $(0,t)$ for $t \in (0,T)$, applying the Cauchy-Schwarz inequality, and using $w(0) = 0$ yields

$$(6.27) \quad \|w(t)\|_E^2 - \frac{\lambda_1}{2} \|w(t)\|_0^2 \le \frac{1}{2} \int_0^t e^{\lambda_1 s} \|e^1_s(s)\|_0^2 ds,$$

which by the definitions of $w, \lambda_1$, and $\hat{C}$ becomes

$$(6.28) \quad \|e^{21}(t)\|_E \le \{\int_0^t e^{-\hat{C}(t-s)} \|e^1_s(s)\|_0^2 ds\}^{1/2}.$$

By (6.18), $B(e^1_s(s),\psi) = 0 \quad \forall s,\psi \in (0,t) \times S(\Delta)$, and so by a standard duality argument we have that

$$(6.29)\qquad \|e^{21}(t)\|_E \le Ch(\Delta)\{\int_0^t e^{-\hat{C}(t-s)}\|e_s^1(s)\|_E^2 ds\}^{1/2}.$$

Using the result in [2, Thm. 7.1], as in the demonstration of (6.22), we get (6.23) from (6.29).

Defining now $z(s) = e^{\lambda_1 s/2} e^{22}(s)$, from eqs. (6.21) we get

$$(6.30)\qquad \begin{cases} \langle z_s(s),\phi\rangle + B(z(s),\phi) - \frac{\lambda_1}{2}\langle z(s),\phi\rangle = 0; \\ \qquad\qquad \forall s,\phi \in (0,T) \times H_0^1, \end{cases}$$

$$z(0) = e^2(0).$$

Setting $\phi = 2z_s(s)$ in (6.30) and using the same arguments which led to (6.28) yields

$$(6.31)\qquad \|e^{22}(t)\|_E \le 2^{1/2}e^{-\hat{C}t/2}\|e^2(0)\|_E.$$

Consecutively setting $\phi = z(s)$ and $2z_s(s)$ in (6.30) and integrating on $(0,t)$ yields

$$(6.32)\qquad 2t\{\|z(t)\|_E^2 - \frac{\lambda_1}{2}\|z(t)\|_0^2\} \le \|z(0)\|_0^2,$$

which by the definitions of $z$, $\lambda_1$, and $\hat{C}$ becomes

$$(6.33)\qquad \|e^{22}(t)\|_E \le \frac{e^{-\hat{C}t/2}}{t^{1/2}}\|e^2(0)\|_0, \quad \text{for } t > 0.$$

It follows from (6.3), (6.14), and (6.15) that

$$(6.34) \qquad B(e^2(0),\psi) = 0 \quad \forall\psi \in S(\Delta),$$

which again by a standard duality argument shows that

$$(6.35) \qquad \|e^2(0)\|_0 \le Ch(\Delta)\|e^2(0)\|_E,$$

and by the result in [2, Thm. 7.1] that

$$(6.36) \qquad \|e^2(0)\|_E \le CE(0,Le^2(0)-0,0)$$

$$= CE(0,U_t^\Delta(0)+Lu_0-f(0),0).$$

Combining (6.31), (6.33), (6.35), and (6.36), we have (6.24), thus completing the proof of Thm. 6.1.

## 7. Computational Procedures and Examples

In this section we outline the main features of a general FEMOL program which includes a posteriori error analysis. The program is of a research type for assessing various aspects and performances of the FEMOL procedure. While it is not presently for commercial use, it includes many user-oriented features employed in available commercial software. The experiments using the program were directed toward the evaluation of the efficiency of the approach and the applicability of conclusions based upon the asymptotic analysis presented in Sections 4-6.

The discussion here is principally oriented to selected examples, in which the exact solutions are available. These solutions have forms typical of those arising often with systems of linear and nonlinear parabolic equations used in various applications, and include decay and sharp transitions in time, oscillations, and travelling waves.

The model problem considered is that given by eqs. (1.1), where the coefficients $a(x) = \cosh(4x-2)$ and $b(x) = \sinh(2x)$ are as pictured in figure 2, and the functions $u_0$, $f$, and $g$ are chosen to be the smooth, compatible data such that the exact solution $u$ of eqs. (1.1) is as given in each of the examples below. The theory in sections 2-6 was presented for the case when the boundary data $g \equiv 0$, but holds more generally for eqs. (1.1).

For $t > 0$, our goal is estimate $|||e(t,\cdot)|||$ with $\mathcal{E}(t)$, where $e = u - U$ is the space discretization error, the computable a posteriori estimator $\mathcal{E}(\cdot)$ is as defined in (1.4) and (5.9), and the norm $|||\cdot|||$ is as defined in (1.2). The theory in sections 2-6 was presented for the estimation of $\|e(t,\cdot)\|_E$, where the norm $\|\cdot\|_E$ is defined in (2.4), but we note that up to higher order terms in the space mesh size, $\|e(t,\cdot)\|_E \sim |||e(t,\cdot)|||$.

Our evaluation of $\mathcal{E}(\cdot)$ is based upon the behavior of the effectivity ratio $\Theta(\cdot)$, defined in (1.5), and we shall examine how it depends on the relative error $E_{REL}(\cdot)$, as defined in (1.9).

The theory given in Section 5 showed that for sufficiently (very) fine partitions of the interval $(0,1)$, $1/\Theta(t)$ is bounded uniformly for all $t$ in a time interval of interest, and $\Theta(t) \to 1$ as the partitions are refined. We shall see that $\Theta(t)$ is near 1 for practically <u>any</u> partition of $(0,1)$ in the examples, and in fact appears to converge very rapidly to 1 as the partitions are refined.

Before presenting our results, we first discuss some of the relevant aspects of the computational procedures used in the experiments and more general problems. For further details we refer to a following paper [7] and [8].

All computations were carried out in double precision arithmetic by the research program FEMOL1, which was written and implemented on the IBM 370 System in the Division of Computer Pesearch and Technology, NIH. FEMOL1 has the capabilities to solve coupled systems of linear and nonlinear parabolic PDEs with piecewise linear elements in one space dimension, with smooth or unsmooth initial data, and general separated end-point boundary conditions. The program has four basic stages of operation, consisting of (1) the processing of control information, (2) the assembly of matrices, (3) the initialization of data, and (4) the time-integration of the resulting ODE system. Many of the standard user-oriented features and logical switches employed in these four stages were borrowed from existing software, much of which was surveyed in [16]. When discontinuous changes in the space mesh are allowed, decisions concerning these changes are made adaptively and looping occurs in stages 2-4. We shall restrict most of the discussion to those features relevant to the experiments conducted here.

In the control stage, flags are set which govern the number and type of PDEs solved, the space mesh used, and whether or not the error is to be estimated, the true error is to be computed, or adaptive mesh construction is to be employed. Here, the target time points for output are also determined. In the examples to be presented, 100 equally spaced time points $\{T_m\}_{m=1,100}$ were distributed over the time intervals of interest, at which $\mathcal{E}$, $|||e|||$, $E_{REL}$, and $\Theta$ were computed, but other selections for output are possible.

As in the computation of $\mathcal{E}$, $|||e|||$, $E_{REL}$, and $\Theta$, FEMOL1 uses standard two-point Gaussian quadrature in the assembly stage to compute the mass and stiffness matrices and load vector needed for the reduction of eqs. (1.1) to an implicit system of ODEs. This process of course introduces the effects of the numerical integration into the computations. After minimal testing, this effect was accepted for our experiments with time-independent space meshes. In some of the computations performed with adaptively constructed meshes, however, these pollution effects were more significant and an alternate procedure had to be adopted (cf. a following paper [7]).

With the smooth, compatible data in the examples here, the discrete initial data was determined in the initialization stage as the standard finite element approximation of the solution of an elliptic boundary value problem. When the initial data is not smooth, FEMOL1 chooses the linear interpolation as the discrete approximation.

The time-integration of the resulting ODE system is accomplished in a subroutine with a version of Gear's variable-step, variable-order backward differentiation formulas which we modified to efficiently handle matrices with banded structure. The ODE solver is essentially equivalent to that in [12]. To ensure that the effects of the time discretization did not pollute the results, very stringent error requirements were placed on the time discretization scheme. The ODE solver used places equal weights on the $N-1$ components $\{U(t,x_j)\}_{j=1,N-1}$ in attempting to control the relative time discretization error per step. These are not the best criteria, in that each component is given equal weight and that an error per unit step tolerance probably has more meaning. While some minor modifications in this procedure were made, these inefficiencies were accepted for the sake of illustrating the theoretical results. We experimentally adjusted the input tolerance until the discrete time derivative $U_t$ in each example changed by less than $10^{-3}$ throughout the time intervals of interest. This required relative error per step tolerances on the order of $10^{-7}$ to $10^{-8}$. We emphasize that these stringent requirements should not be used in practice, but were made here to isolate the effects to be studied.

One of the main purposes in studying a posteriori error estimates in the FEMOL solution of PDEs is the development of more efficient algorithms. However, because of the stringent requirements set above and the mentioned, and other not mentioned inefficiences which were accepted for our convenience, a quotation of CPU times used in the examples is not relevant.

As a final computational note, we remark that all of the logical switches and capabilities of FEMOL1 can be set for a given problem by supplying subroutines for the problem data and modifying less than 20 cards of input data.

Example 1. For $0 \le t \le 3$ the exact solution $u$ of eqs. (1.1) is taken to be

$$(7.1) \quad u(t,x) = u^1(t,x) + [\tfrac{1}{2} + \tfrac{1}{2}\tanh(10t-11)] \cdot u^2(t,x),$$

where

$$(7.2) \quad u^1(t,x) = 1 - \exp\left\{\frac{-x^2}{(10t+.1)}\right\}, \text{ and}$$

$$(7.3) \quad u^2(t,x) = 2\sin(\pi x) + 2\cos(2\pi t)\sin(2\pi x).$$

For $0 \le t \le .9$ $u^1$ completely dominates the behavior of $u$ and $u$ decays in $t$ (cf. figure 3(a)). For $.9 < t < 1.1$ $u$ undergoes a sharp transition, and for $1.1 \le t \le 3$ the oscillatory character of $u^2$ is dominant (cf. figure 3(b)).

The FEMOL was implemented in this example for a sequence of three uniform space meshes, using 10, 20, and 40 finite elements. Figure 4 shows the behavior of the errors in time, and the linear rate of convergence is apparent. For $t > 1$ the errors in this sequence oscillate about the 10%, 5%, and 2.5% levels. Figure 5 illustrates how well the estimator $E(\cdot)$ works. For any of the meshes and all $t \in [0,3]$ we see that the effectivity ratio $\Theta(\cdot)$ satisfies

(7.4) $$|\Theta(t) - 1| \leq .07,$$

and that with 40 elements

(7.5) $$|\Theta(t) - 1| \leq .005; \ \forall t \in [0,3].$$

In addition to these unexpectedly good results, it appears that

(7.6) $$\Theta(t) = 1 + O(N^{-2}) \quad \text{as} \quad N \to \infty; \quad \forall t \in [0,3],$$

a result much better than that which we could prove in Thm. 5.2.

Example 2. For $0 \leq t \leq .1$ the exact solution of eqs. (1.1) is taken to be that defined in (1.8), namely

(7.7) $$u(t,x) = \frac{1}{2} + \frac{1}{2} \tanh[2\beta(x-10t)]; \quad \beta = 20.$$

The solution is a wave with approximate front width $\beta^{-1}$ which moves in the positive x-direction at speed 10, and is as pictured in *figure 1 in section 1*.

As in Example 1, the estimator $E(\cdot)$ was tested by implementing the FEMOL for a sequence of uniform space meshes, using here 20, 40, 80, and 160 finite elements. In figure 6 we see the oscillatory temporal character of the error for 20 and 40 elements, due to the inability of the mesh to resolve the wave front. The error again appears to converge linearly for all $t$, provided that $N \geq 40$. Note that for $N = 20(40)$ the error oscillates about the 50% (25%) mark. In figure 7 we see that even in the case of 50% error, the effectivity ratio $\Theta(\cdot)$ satisfies

$$(7.8) \qquad \frac{1}{2} \leq \Theta(t) \leq 2; \quad \forall t \in [0,.1].$$

Figure 8 illustrates that with an average error of about 25%, for almost all $t \in [0,.1]$

$$(7.9) \qquad |\Theta(t)-1| \leq .1,$$

and that when $N = 160$ and the error is about 6% for each $t$, we have

$$(7.10) \qquad |\Theta(t)-1| \leq .005; \quad \forall t \in [0,.1].$$

The quality of these results does not appear to be limited to the use of uniform space meshes, as experiments with slightly nonuniform meshes in the same examples has indicated. Also, we shall see in a following paper [7] that the adaptively constructed meshes for Example 2 can be very nonuniform, and yet the analogous effectivity ratio is close to 1.

We finally remark that the norms $\|\cdot\|_E$ and $|||\cdot|||$ used in the analysis and computations were chosen because they are the natural norms arising from the weak form of eqs. (1.1). In applications where there is no natural energy associated with the system, or when some other characteristic is of primary interest, then it may be desirable to estimate a different norm of the error. By altering the approximating finite element subspaces and a posteriori estimator in our procedure appropriately, such estimates can be obtained. Nonetheless, the above results illustrate the robustness of the estimator $E(\cdot)$, even when apparently out of the asymptotic range.

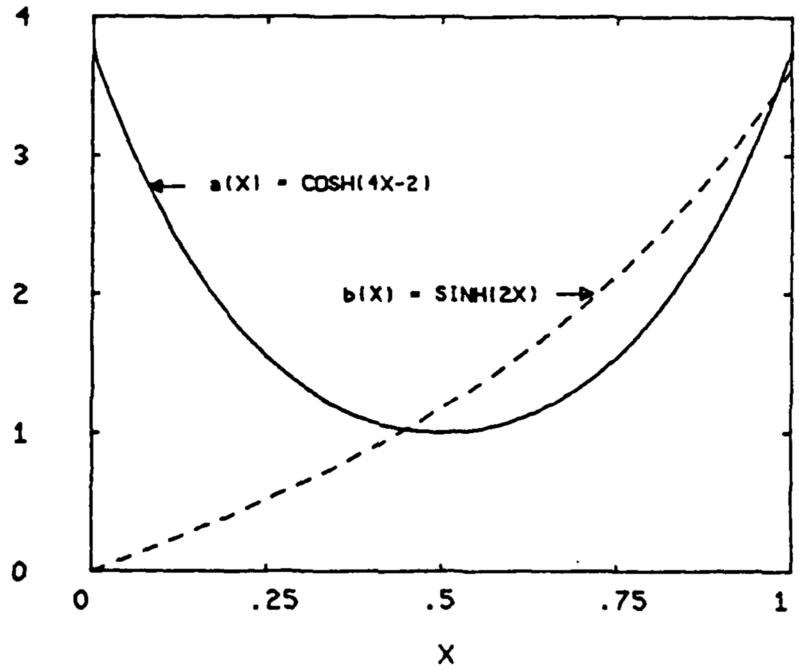


Figure 2. Coefficients of Eqs. (1.1) vs. X

(a)

(b)

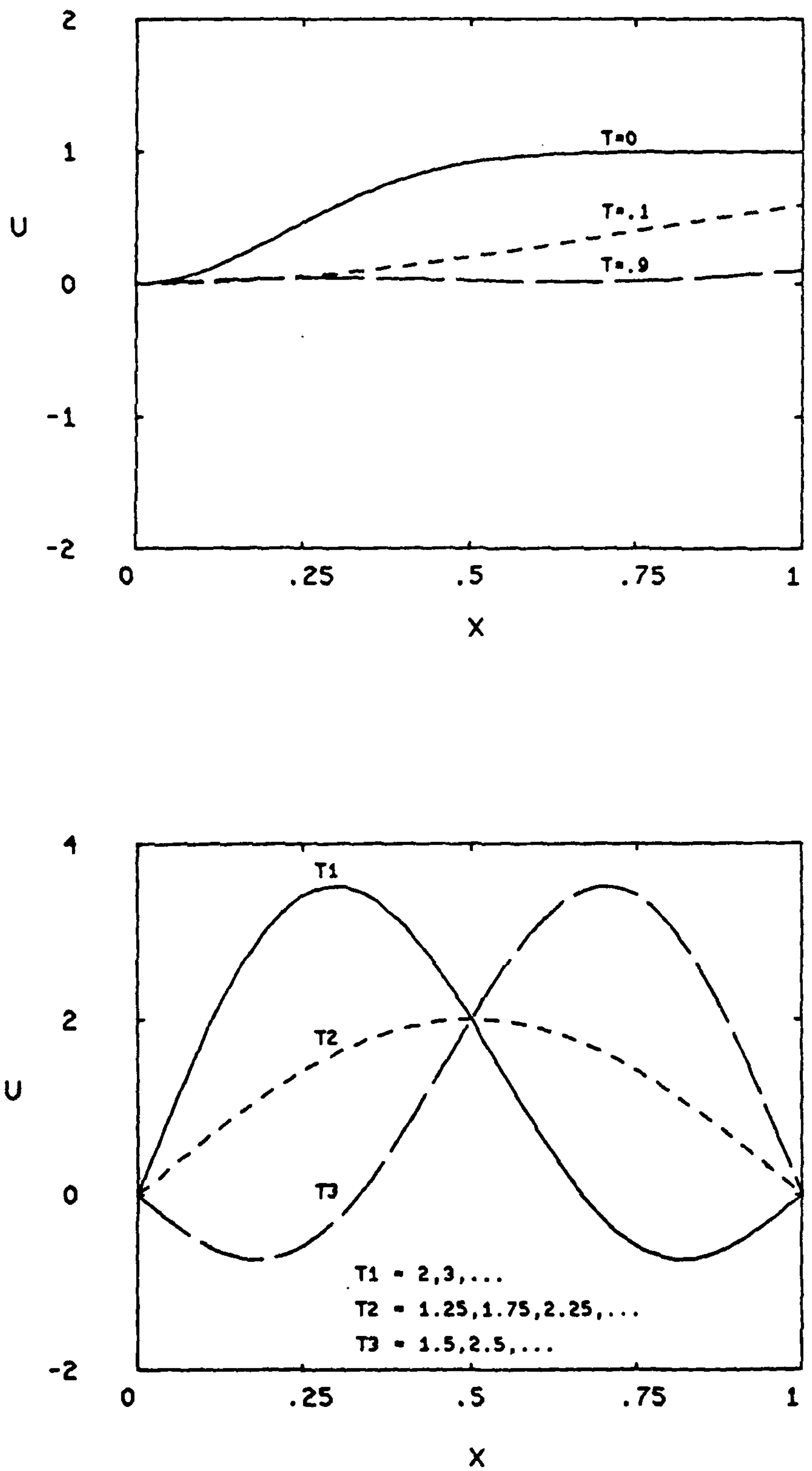


Figure 3. Exact Solution u(T,X) of Example 1 vs. X for Various T.

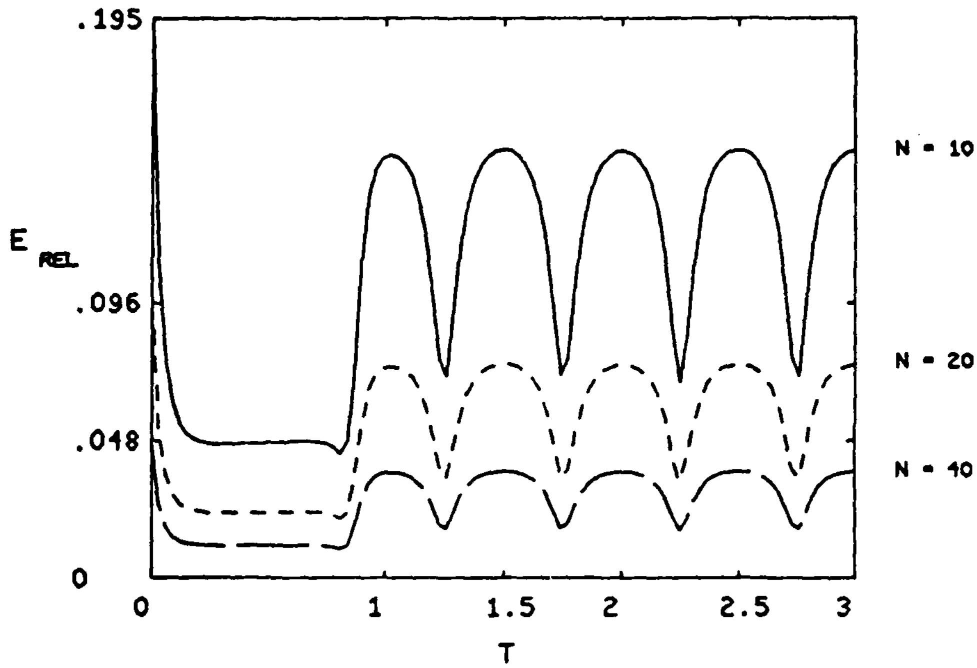


Figure 4. Relative Error $E_{REL}(T)$ in Example 1 vs. T. N = Number of Uniform Finite Elements

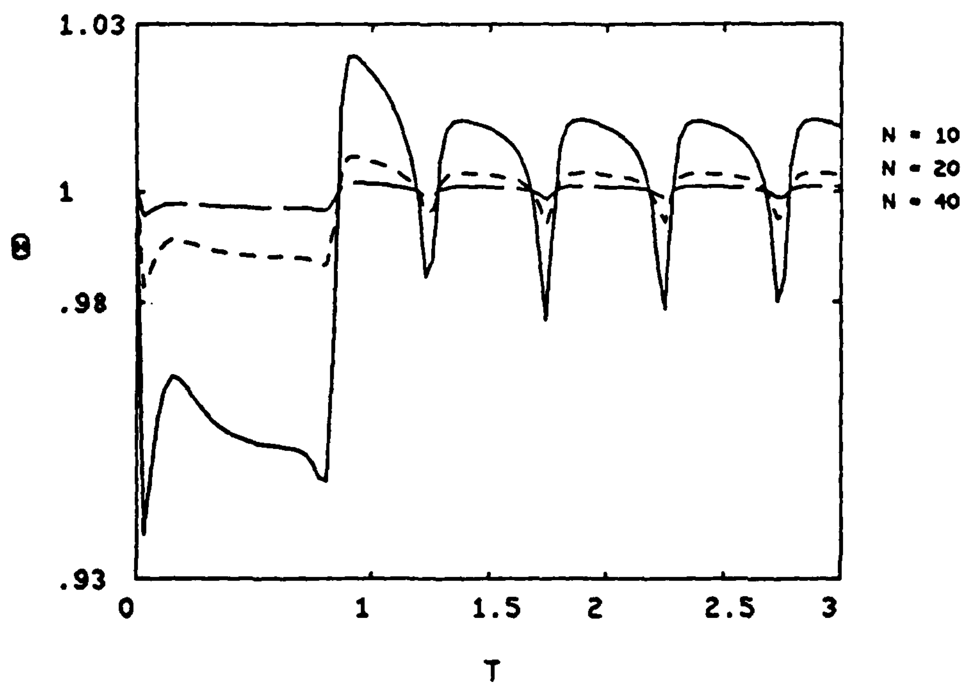


Figure 5. Effectivity Ratio $\Theta(T)$ in Example 1 vs. T. N = Number of Uniform Finite Elements.

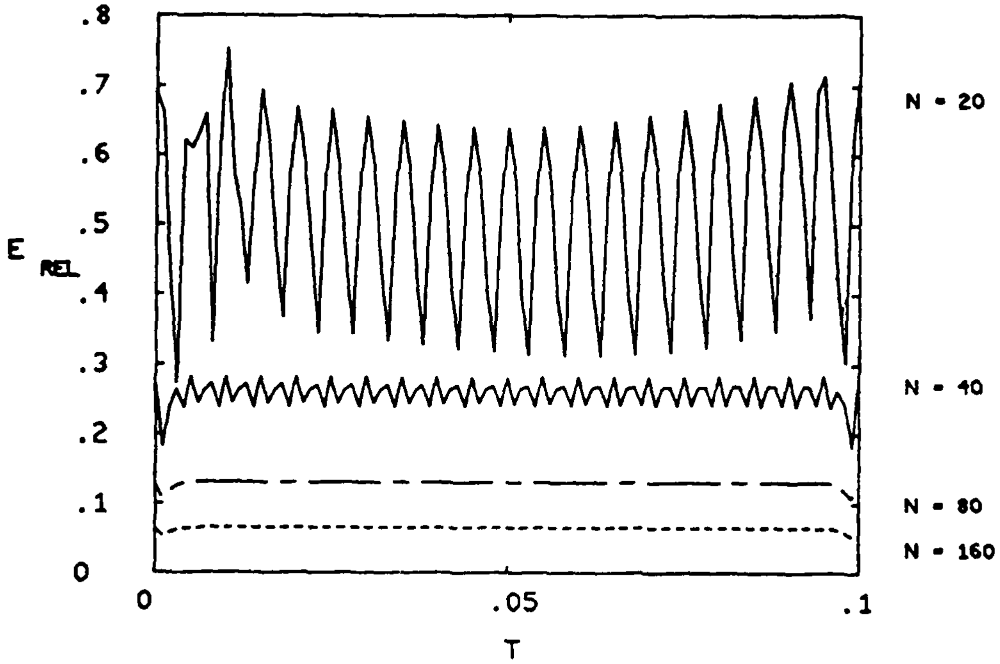


Figure 6. Relative Error $E_{REL}(T)$ in Example 2 vs. T.
N = Number of Uniform Finite Elements.

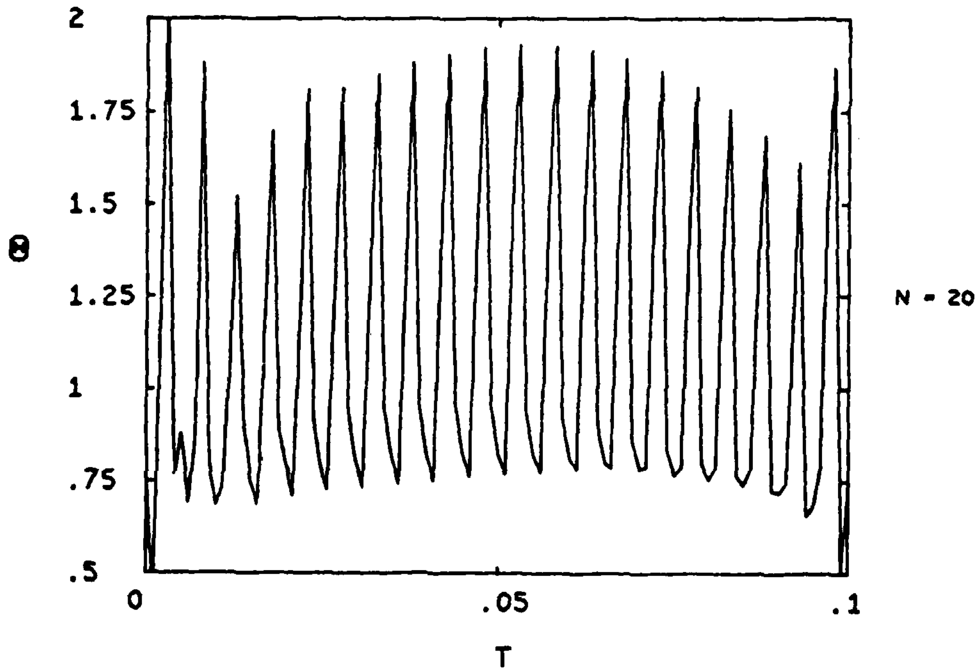


Figure 7. Effectivity Ratio $\Theta(T)$ in Example 2 vs. T.
N = Number of Uniform Finite Elements.

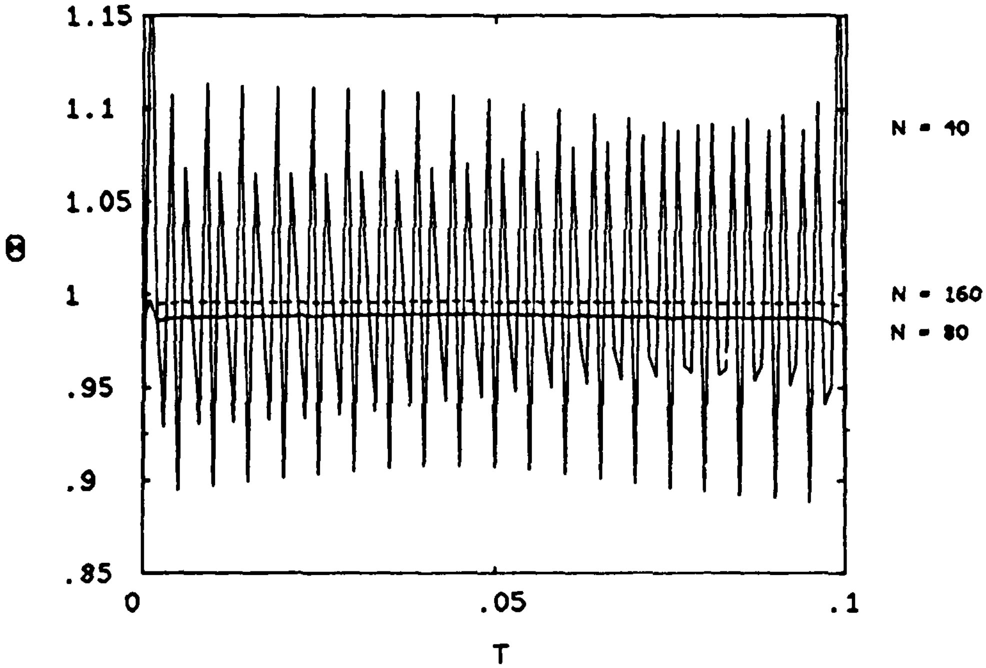


Figure 8. Effectivity Ratio Θ(T) in Example 2 vs. T. *N = Number of Uniform Finite Elements.*

## References


1. Babuška, I., Luskin, M.: An adaptive time discretization procedure for parabolic problems. In: Advances in comp. meth. for partial differential equations IV (R. Vichnevetsky, R.S. Stepleman, eds.) 18, pp. 5-13 (1981).

2. Babuška, I., Rheinboldt, W.: A posteriori error analysis of finite element solutions of one-dimensional problems. SIAM J. Numer. Anal. 18, pp. 565-589 (1981).

3. Babuška, I., Rheinboldt, W.: Analysis of optimal finite element meshes in $\mathbb{R}^1$. Math. Comp. 33, pp. 435-463 (1979).

4. Babuška, I., Rheinboldt, W.: A posteriori error estimates for the finite element method. Int. J. Numer. Methods Eng. 12, pp. 1597-1615 (1978).

5. Babuška, I., Rheinboldt, W.: Error estimates for adaptive finite element computations. SIAM J. Numer. Anal. 15, pp. 736-754 (1978).

6. Babuška, I., Aziz, A.K.: Survey lectures on the mathematical foundations of the finite element method. In: The mathematical foundations of the finite element method with applications to partial differential equations (A.K. Aziz, ed.), pp. 1-359. New York: Academic Press 1972.

7. Bieterman, M., Babuška, I.: The finite element method for parabolic equations, II. A posteriori error estimation and adaptive approach. University of Maryland I.P.S.T. Technical Note BN-984, 1982.

8. Bieterman, M.: Ph.D. Thesis, University of Maryland, 1982.

9. Bramble, J.H., Schatz, A.H., Thomee, V., Wahlbin, L.B.: Some convergence estimates for Galerkin type approximations for parabolic equations. SIAM J. Numer. Anal. 14, pp. 218-241 (1977).

10. Douglas, J. Jr., Dupont, T., Wheeler, M.F.: A quasi-projection analysis of Galerkin methods for parabolic and hyperbolic equations. Math. Comp. 32, pp. 345-362 (1978).

11. Grisvard, P.: Characterisation de quelques espaces d'interpolation. Arch. Rational Mech. Anal. 25, pp. 40-63 (1967).

12. Hindmarsh, A.: Preliminary documentation of GEARIB: solution of implicit systems of ordinary differential equations with banded Jacobians. Report UCID-30130, Lawrence Livermore Lab., Livermore, California, 1976.

13. Kato, T.: Perturbation theory for linear operators. New York: Springer-Verlag 1966.

14. Krein, S.G.: Linear differential equations in Banach space. AMS Translations of Mathematical Monographs 29, 1971.

15. Lions, J.L., Magenes, E.: Nonhomogeneous boundary value problems and applications I and II. New York: Springer-Verlag 1973.

16. Machura, M., Sweet, R.: A survey of software for partial differential equations. ACM-TOMS 6, pp. 461-488 (1980).

17. Thomee, V.: Negative norm estimates and superconvergence in Galerkin methods for parabolic problems. Math. Comp. 34, pp. 93-113 (1980).

18. Wheeler, M.F.: A priori $L_2$ error estimates for Galerkin approximations to parabolic partial differential equations. SIAM J. Numer. Anal. 10, pp. 723-759 (1973).

FILME

8-8